

Universidade Federal de Santa Catarina
Programa de Pós-Graduação em Ciência da Computação

Marco Antonio Floriano de Oliveira

Correlacionamento Estéreo de Complexidade Linear
Baseado em Indexação de Regiões

Dissertação submetida à Universidade Federal de Santa Catarina como parte dos requisitos para a obtenção do grau de Mestre em Ciência da Computação

Orientador: Prof. Raul Sidnei Wazlawick, Dr.

Florianópolis, novembro de 2006

Correlacionamento Estéreo de Complexidade Linear Baseado em Indexação de Regiões

Marco Antonio Floriano de Oliveira

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, Área de Concentração em Sistemas de Conhecimento, e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.

Prof. Raul Sidnei Wazlawick, Dr.
Coordenador do curso

Banca examinadora:

Prof. Raul Sidnei Wazlawick, Dr.
Orientador

Prof. Carla Maria dal Sasso Freitas, Dr.

Prof. Mauro Roisenberg, Dr.

Prof. Fernando Álvaro Ostuni Gauthier, Dr.

*The reasonable man adapts himself to the world;
the unreasonable one persists in trying to adapt the
world to himself. Therefore, all progress depends
on the unreasonable man.*

George Bernard Shaw

Àqueles que estão sempre a nosso lado, mesmo
que nem sempre percebamos, nos ajudando a cada
passo da caminhada, com seu amor, carinho,
conhecimento e inspiração.

Sumário

Lista de Figuras.....	vii
Lista de Tabelas.....	x
Resumo.....	xi
Abstract.....	xii
1 Introdução.....	1
1.1 Objetivo.....	3
1.2 Objetivos Específicos.....	3
1.3 Justificativa.....	3
1.4 Estrutura da Dissertação.....	4
2 Revisão Bibliográfica.....	5
2.1 Correspondência Estéreo.....	5
2.1.1 Profundidade e Disparidade.....	6
2.1.2 Retificação do Par de Imagens.....	7
2.2 Restrições Fundamentais.....	9
2.3 Trabalhos Relacionados.....	12
2.3.1 Métodos Globais.....	12
2.3.2 Métodos Locais.....	13
2.3.3 Relação com Métodos Atuais.....	15
3 Definição do Método.....	16
3.1 Indexação de Regiões.....	16
3.1.1 Função de Indexação.....	20
3.1.2 Indexação Baseada em Intensidades.....	21
3.2 Detecção de Falsas Correlações.....	24
3.3 Interpolação do Mapa de Disparidades.....	28
4 Resultados.....	31
4.1 Parametrização.....	34
4.2 Metodologia de Teste.....	36
4.3 Conjuntos de Dados Padrão.....	36
4.3.1 Comparativo entre Métodos.....	41

5 Conclusões.....	43
5.1 Trabalhos Futuros.....	44
5.1.1 Refinamentos.....	44
5.1.2 Extensão.....	45
Referências Bibliográficas.....	46
Anexo 1 – Conjuntos de Dados.....	49
Anexo 2 – Plataforma Computacional.....	50
Glossário.....	51

Lista de Figuras

Figura 1. Par estéreo e mapa com a disparidade para cada ponto na imagem esquerda em relação a seu correspondente na direita.....	1
Figura 2. Relação entre profundidade e disparidade para planos de projeção coplanares.....	6
Figura 3. Geometria epipolar para um sistema de duas câmeras.....	7
Figura 4. Espaço de correlação para planos de projeção coplanares.....	7
Figura 5. Ambigüidade na correspondência estéreo.....	9
Figura 6. Restrição de unicidade: (a) exemplo; (b) exceção (correlacionamento em áreas de semi-oclusão).....	9
Figura 7. Restrição de continuidade: exemplo (superfície do mesmo objeto); exceção (bordas entre objetos a distâncias diferentes).....	10
Figura 8. Restrição de similaridade: semelhança fotométrica entre pontos correspondentes (áreas escuras); exceção (pontos mais claros).....	10
Figura 9. Restrição de ordenação: exemplo (pontos A, B, C, D); exceção (ponto P, distância relativa muito grande entre os objetos).....	11
Figura 10. Restrição de ordenação: função de disparidade monótona.....	11
Figura 11. Medida de similaridade para um único ponto ao longo da linha epipolar correspondente (correlação x disparidade).....	14
Figura 12. Indexação de regiões: (a) duas correlações; (b) limitação da correlação (disparidade positiva).....	17
Figura 13. Algoritmo básico para indexação de regiões.....	18
Figura 14. Problema de repetição do discriminante: (a) falsa correlação; (b) deslocamento horizontal.....	18
Figura 15. Modificação do algoritmo para indexação em uma linha deslocada.....	19
Figura 16. Indisponibilidade do índice.....	19
Figura 17. Exemplo de cálculo do índice de região.....	22
Figura 18. Algoritmo para aplicação da restrição de continuidade.....	26

Figura 19. Cálculo do vetor de pesos W a partir da suavização do histograma do mapa de entrada (a esquerda) e cálculo do histograma ponderado U para a primeira janela (a direita).....	26
Figura 20. Aplicação da restrição de continuidade para um ponto (i, j)	27
Figura 21. Atualização do histograma para janelas em $(i, j+1)$, $(i, j-1)$ e $(i+1, j)$, respectivamente.....	27
Figura 22. Linhas e colunas usadas na interpolação com base no ponto mais próximo.....	28
Figura 23. Exemplo de interpolação com base no ponto mais próximo (disparidades originais à esquerda e em azul).....	28
Figura 24. Algoritmo de interpolação: inicialização (à esquerda); obtenção da disparidade para um ponto (i, j) (à direita).....	29
Figura 25. Algoritmo de interpolação: geração do mapa intermediário.....	29
Figura 26. Algoritmo de interpolação: geração do mapa final.....	30
Figura 27. Resultado da indexação de regiões: (a) imagem esquerda; (b) disparidades calculadas (indexações: 51%, correlações: 27%);.....	31
Figura 28. Mapas de eficiência da indexação de regiões: (a) colisões (em cinza) e correlações (em branco); (b) disparidades negativas (rejeitadas durante a indexação);.....	32
Figura 29. Histograma de ocupação dos vetores de indexação para cada linha epipolar (cada linha da imagem corresponde a um vetor).....	32
Figura 30. Resultado após a restrição de continuidade: (a) para as disparidades calculadas na indexação (densidade: 13%); (b) para todos os pontos no mapa (densidade: 40%).....	33
Figura 31. Interpolação baseada no ponto mais próximo: (a) mapa não equalizado; (b) mapa equalizado.....	33
Figura 32. Número de segmentos (em preto) e variação (em branco) em relação ao tamanho S do índice de segmento (para intensidades de 8 bits).....	35
Figura 33. Resultados no conjunto “tsukuba”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).....	36
Figura 34. Resultados no conjunto “map”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).....	37

Figura 35. Resultados no conjunto “venus”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).....	38
Figura 36. Resultados no conjunto “sawtooth”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).....	38
Figura 37. Resultados no conjunto “cones”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).....	39
Figura 38. Resultados no conjunto “teddy”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).....	40
Figura 39. Tempo de processamento para conjuntos de dados com diferentes resoluções.....	41
Figura 40. Tempos de processamento de diferentes métodos para conjuntos de dados com diferentes resoluções.....	42

Lista de Tabelas

Tabela 1. Parâmetros do algoritmo de indexação de regiões.....	23
Tabela 2. Parâmetros do algoritmo de detecção de falsas correlações.....	25
Tabela 3. Valores padrão para os parâmetros do método.....	34
Tabela 4. Acurácia e tempo de processamento.....	40

Resumo

Este trabalho apresenta um método de complexidade linear para correlacionamento estéreo de tempo-real, no qual o tempo de processamento depende apenas da resolução do par de imagens. Regiões ao longo de cada linha epipolar são indexadas para produzir o mapa de disparidades, ao invés de realizar uma busca pela melhor correlação. Métodos locais atuais não possuem complexidade linear, uma vez que dependem de uma busca através de um espaço de possíveis correlações. O método apresentado é limitado a um conjunto de câmeras paralelas ou um par de imagens retificadas, porque todas as disparidades devem ocorrer na mesma direção e sentido. Uma restrição de continuidade é aplicada para remover falsas correlações. O mapa resultante é semi-denso, mas as disparidades são bem distribuídas e as áreas esparsas são preenchidas satisfatoriamente através de interpolação baseada no ponto mais próximo. Nenhum ajuste específico de parâmetros é necessário. Resultados experimentais com conjuntos de dados padrão alcançam aproximadamente 90% de acurácia, usando metodologia de teste bem conhecida e os mesmos parâmetros em todos os testes.

Palavras-chave: visão estéreo, complexidade linear, tempo-real.

Abstract

This work presents a linear complexity method for real-time stereo matching, in which the processing time is only dependent on the resolution of the image pair. Regions along each epipolar line are indexed to produce the disparity map, instead of searching for the best match. Current local methods have non-linear complexity, as they rely on searching through a correlation space. The present method is limited to a parallel camera setup or a rectified image pair, because all disparities must occur in the same direction and sense. A continuity constraint is applied in order to remove false matches. The resulting map is semi-dense, but disparities are well distributed and sparse areas are satisfactorily filled by nearest interpolation. No specific parameter tuning is required. Experimental results on standard datasets reach around 90% of accuracy, using well-known test methodology and the same parameters in all tests.

Key words: stereo vision, linear complexity, real-time.

1 Introdução

Estabelecer a correspondência entre os pontos de duas imagens tiradas da mesma cena permite que sejam calculadas as distâncias entre o observador e as superfícies dos objetos visualizados. Com o mapeamento de pontos entre uma imagem e outra de um par estéreo, é possível calcular a disparidade relativa para cada par de pontos correspondentes, ou seja, a diferença de posição, em coordenadas da imagem, entre projeções de um mesmo ponto no espaço. A partir dessas disparidades, é possível estimar a distância entre o par de imagens (observador) e pontos observados na cena (superfícies dos objetos visualizados). [6]

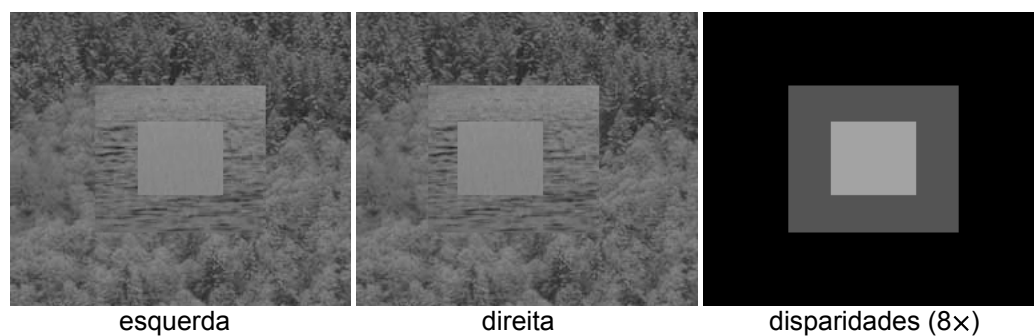


Figura 1. Par estéreo e mapa com a disparidade para cada ponto na imagem esquerda em relação a seu correspondente na direita.

O problema da correspondência estéreo é o cerne da visão estéreo, que é uma área com aplicação em sistemas de reconstrução 3D, navegação autônoma, reconhecimento de objetos, monitoramento automático, segmentação de imagem etc. Uma solução de baixa complexidade e razoável acurácia para esse problema é bastante interessante, uma vez que muitos desses sistemas são intrinsecamente de tempo-real.

A visão estéreo pode ser ativa, na qual o observador age sobre o ambiente através de emissões de padrões luminosos (como luz estruturada ou varredura a laser), ou passiva, na qual as imagens são obtidas sem influenciar a cena observada. Na visão passiva, o problema da correspondência deve ser resolvido apenas com base na informação contida nas imagens, não havendo qualquer controle sobre o ambiente. [9]

Métodos de visão estéreo passiva destinados a sistemas de tempo-real diferem quanto a restrições, otimizações e técnicas para tentar reduzir o espaço de busca pela melhor correlação [11, 16, 17, 37]. Entretanto, a abordagem geral utilizada até então por esses métodos é utilizar uma determinada função de correlação para comparar cada ponto em uma imagem com um certo conjunto de pontos não unitário na outra (quer sejam avaliados valores de intensidade luminosa ou atributos de aspectos extraídos das imagens) [35]. Embora esses métodos possuam baixa complexidade comparados àqueles baseados em minimização de energia (função de custo global) [4, 13, 20], eles ainda apresentam complexidade não-linear em relação ao número de pontos da imagem, por causa da natureza combinatória do processo de busca pela melhor correlação dentre os candidatos na outra imagem para cada ponto.

Este trabalho apresenta um método de complexidade linear para o problema da correspondência estéreo, no qual o tempo de processamento depende apenas da resolução do par de imagens e a acurácia é de aproximadamente 90%. O método é composto de uma seqüência de três algoritmos para geração do mapa de disparidades: o primeiro para indexação de regiões e cálculo inicial das disparidades, o segundo para detecção de falsas correlações por meio de uma restrição de continuidade e o último para rápida interpolação do mapa de disparidades resultante. [32]

A principal contribuição deste trabalho é um processo de indexação de regiões para realizar o correlacionamento estéreo, evitando a natureza combinatória de uma busca pela melhor correlação para cada ponto. Isso permite que cada ponto seja processado em tempo constante, resultando em complexidade linear para processar todo o par de imagens. Um método com essa característica é desejável em sistemas de tempo-real que necessitem de visão estéreo.

Como todas as disparidades devem ocorrer na mesma direção e sentido (os planos de projeção devem ser coplanares), o método é limitado a um conjunto de câmeras paralelas ou um par de imagens retificado (reprojeção das imagens com rotação dos planos em relação aos centros de projeção). Depois de aplicada a restrição de continuidade e feita a interpolação, o mapa resultante apresenta a maior parte do erro em áreas de descontinuidade, ou seja, nas bordas dos objetos.

1.1 Objetivo

Esta pesquisa foi realizada com o objetivo de formular um método de visão estéreo (cálculo do mapa de disparidades para um par estéreo) com acurácia semelhante àquela encontrada em métodos existentes, mas com a vantagem de apresentar complexidade linear em relação ao número de pontos, independente do conteúdo da imagem (processamento de cada ponto em tempo constante).

1.2 Objetivos Específicos

Os objetivos específicos deste trabalho são:

- (a) Estabelecer a correspondência entre regiões correlatas sem usar uma função para calcular o grau de similaridade entre elas, isto é, sem compará-las diretamente;
- (b) Encontrar uma forma eficiente de remoção de falsas correlações no mapa de disparidades resultante do processo de indexação de regiões que também possua complexidade linear;
- (c) Dispor de uma interpolação para geração do mapa final que, além da mesma complexidade linear das fases anteriores, tenha um impacto proporcionalmente menor no tempo de processamento total do método;
- (d) Obter resultados (mapas de disparidades densos) com acurácia mínima de 90% a partir de conjuntos de dados padrão (bem conhecidos e aceitos pela comunidade) sem a necessidade de ajustar parâmetros.

1.3 Justificativa

Sistemas nos quais o problema da correspondência estéreo precisa ser resolvido em tempo-real demandam uma solução algorítmica de complexidade baixa e, de preferência, não relacionada ao conteúdo das imagens. Métodos atuais destinados a sistemas de tempo-real [35] ainda têm sua complexidade atrelada de certa forma ao espaço de disparidade, uma vez que não descartam a busca pela melhor correlação.

1.4 Estrutura da Dissertação

No capítulo 2, é abordado o problema da correspondência estéreo e são discutidas as abordagens conhecidas para sua solução. No capítulo 3, é feita uma descrição completa de cada algoritmo do método objeto deste trabalho, bem como uma análise de suas limitações. No capítulo 4, é descrita a metodologia de teste utilizada e são apresentados e discutidos vários resultados, incluindo aqueles produzidos a partir de conjuntos de dados padrão. Por fim, no capítulo 5, são colocadas as conclusões. Detalhes sobre os conjuntos de dados padrão podem ser encontrados no Anexo 1.

2 Revisão Bibliográfica

Estereoscopia, ou visão estéreo, é a estimativa de um mapa de profundidades para um ambiente observado a partir de dois pontos de vista levemente separados. O conhecimento das distâncias entre o observador e as superfícies dos objetos a sua frente (2,5D) pode ser utilizado, por exemplo, na reconstrução 3D do ambiente, no reconhecimento de objetos e na localização e movimentação do observador. [25]

Um mapa de profundidades pode ser obtido a partir das disparidades entre pontos correspondentes em ambas as imagens, uma vez que cada profundidade é inversamente proporcional à distância relativa entre as projeções, em ambas as retinas, de um mesmo ponto no espaço. Assim, a estereoscopia depende da solução do problema da correspondência estéreo, o qual resume-se a encontrar esses pares de pontos e calcular o mapa de disparidades. [26, 33]

2.1 Correspondência Estéreo

Uma imagem é um conjunto de intensidades luminosas que correspondem a projeções em seu plano e não contêm explicitamente informação de profundidade. Também não é possível realizar um mapeamento exato entre pontos correspondentes em um par estéreo, por causa de ambigüidade e oclusão. Assim como a correspondência estéreo, vários problemas relacionados à visão são considerados mal-colocados¹, uma vez que a informação percebida não é suficiente para que seja obtida uma solução única e exata para qualquer entrada. Isso se dá porque sua solução está na inversão do processo que originou as imagens, o qual possivelmente ocorreu com perda de informação. [34]

O problema específico da correspondência estéreo é abordado de forma a obter uma solução satisfatória quanto à acurácia (calcular as disparidades entre pontos supostamente correspondentes com uma margem de erro aceitável) e ao tempo de processamento. Por isso, testes utilizando uma metodologia de avaliação padrão são feitos para comparar os resultados de vários métodos sobre um mesmo conjunto de dados. [35]

¹ Um problema mal-colocado, pela definição de Hadamard, é um problema cuja solução não existe, não é única ou não depende continuamente dos dados (instável mediante perturbações). [15]

2.1.1 Profundidade e Disparidade

A relação entre a profundidade de um ponto observado e a disparidade entre suas projeções é inversamente proporcional para qualquer formato e disposição das superfícies de projeção. O modelo no qual as superfícies de projeção são planas é o mais simples e, por isso, mais utilizado. [3]

Se os planos de projeção forem coplanares (câmeras paralelas e ortogonais à linha de base), essa relação pode ser obtida através de uma simples semelhança de triângulos, como é ilustrado na Figura 2. No caso de não haver coplanaridade, é necessário retificar o par estéreo de modo a reprojetar as imagens em um arranjo de planos coplanares, mantendo os mesmos centros de projeção, como é discutido mais adiante.

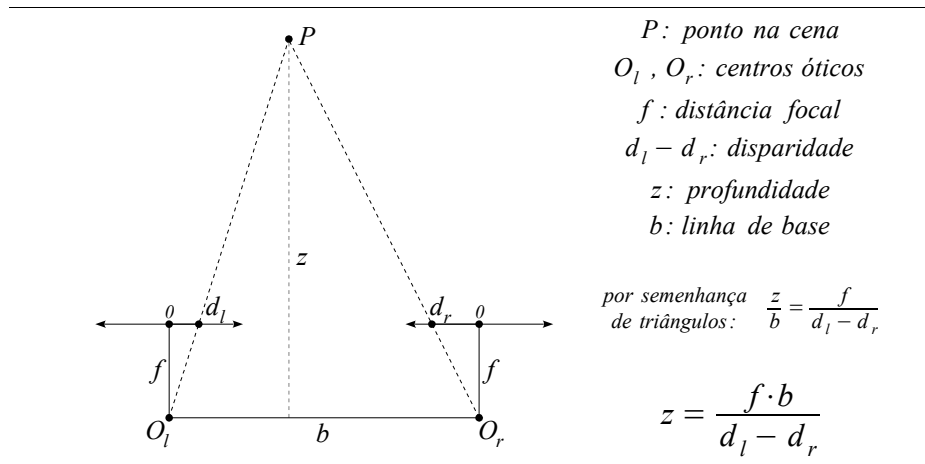


Figura 2. Relação entre profundidade e disparidade para planos de projeção coplanares.

Para calcular a profundidade relativa a uma certa disparidade, é necessário conhecer a distância focal f e o comprimento da linha de base b (distância entre os centros de projeção das duas câmeras), ambos diretamente proporcionais à profundidade z do ponto triangulado. Quanto maior for o comprimento da linha de base, maior é a precisão do cálculo da profundidade, uma vez que a disparidade ocorre com um número maior de pontos. Nesse caso, a diferença entre as imagens também é maior, o que aumenta a quantidade de oclusões e distorções entre áreas correspondentes. Por isso, uma maior precisão implica uma acurácia potencialmente menor. [31]

2.1.2 Retificação do Par de Imagens

O espaço de correlação (conjunto de pontos candidatos a uma correlação) pode ser restrito a apenas uma dimensão, uma vez que as disparidades sempre ocorrem ao longo de linhas epipolares correspondentes, que são as intersecções do plano de triangulação (definido pelos pontos P , O_l e O_r na Figura 3) com os planos de projeção. O par estéreo pode ser retificado de maneira a fazer coincidir as linhas epipolares com as linhas horizontais em ambas as imagens, facilitando computacionalmente seu percorrimento. [24]

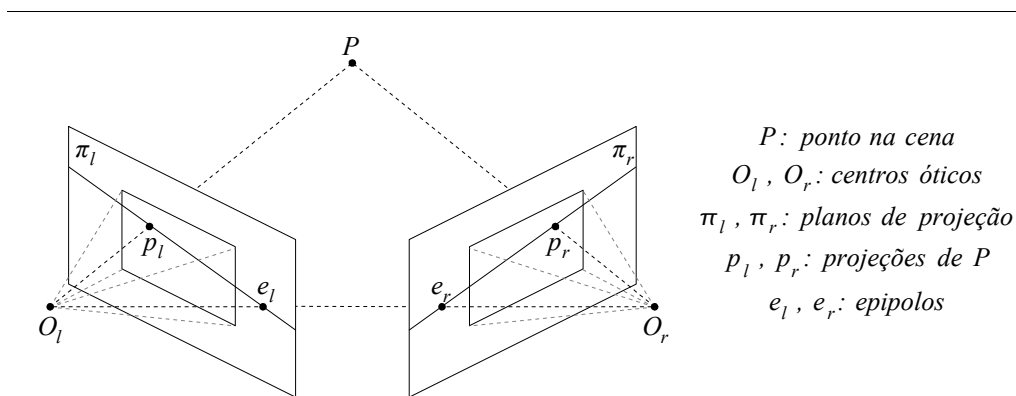


Figura 3. Geometria epipolar para um sistema de duas câmeras.

Uma linha epipolar é a intersecção entre o plano de projeção e o plano que contém a linha de base e o ponto triangulado. O epipolo de um plano de projeção (e_l e e_r na Figura 3) é o ponto de intersecção de todas as linhas epipolares e corresponde à projeção do centro ótico da outra câmera. [1]

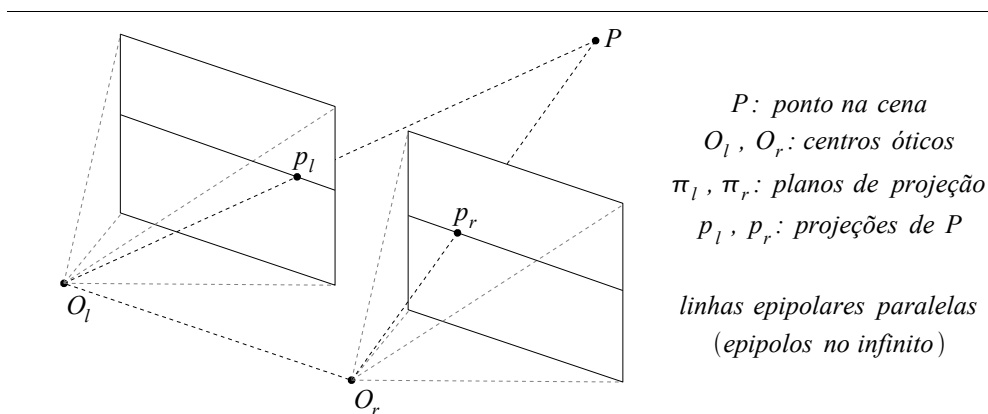


Figura 4. Espaço de correlação para planos de projeção coplanares.

Embora o epipolo pertença ao plano da imagem, ele dificilmente estará dentro da área da imagem, exceto em casos onde a inclinação entre os planos seja bastante acentuada. Com a retificação das imagens (Figura 4), os planos de projeção passam a ser coplanares, o que faz com que cada linha epipolar seja paralela à linha de base e colinear a sua correspondente na outra imagem, deslocando os epipolos para o infinito. [12]

A geometria epipolar depende somente da posição e da orientação relativas entre as câmeras (parâmetros extrínsecos) e do modelo de câmera (parâmetros intrínsecos). Os sistemas de coordenadas das duas câmeras são relacionados através de uma rotação R (orientação relativa) e uma translação T (posição relativa). A representação algébrica da geometria epipolar para um par de câmeras calibradas (ou seja, com parâmetros intrínsecos conhecidos) é a seguinte:

$$p_r^T E p_l = 0 \quad E = R[T]_x \quad T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad [T]_x = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix}$$

Onde E é a matriz essencial [23], que relaciona pontos entre uma imagem outra, expressos no sistema de coordenadas da câmera, com escala indeterminada.

A representação algébrica da geometria epipolar para um par de câmeras não calibradas (ou seja, com parâmetros extrínsecos e intrínsecos desconhecidos) é a seguinte:

$$p_r^T F p_l = 0 \quad F = C_r^{-T} E C_l^{-1} \quad C = \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Onde F é a matriz fundamental [24], que relaciona pontos entre uma imagem e outra, expressos em coordenadas da imagem (em *pixels*), com escala determinada; E é a matriz essencial e C é a matriz de calibração para uma câmera (modelo com distância focal f e posição do centro da imagem).

A matriz fundamental F pode ser estimada diretamente se forem conhecidos alguns pontos correspondentes entre as duas imagens. [10]

2.2 Restrições Fundamentais

O valor de intensidade luminosa de um ponto na imagem não é suficiente para que se possa estabelecer uma correspondência de maneira inequívoca dentro do conjunto de candidatos na outra imagem. Para tentar reduzir essa ambigüidade, foram propostas restrições que limitam o problema da correspondência por meio de suposições que sejam geralmente válidas. [21]

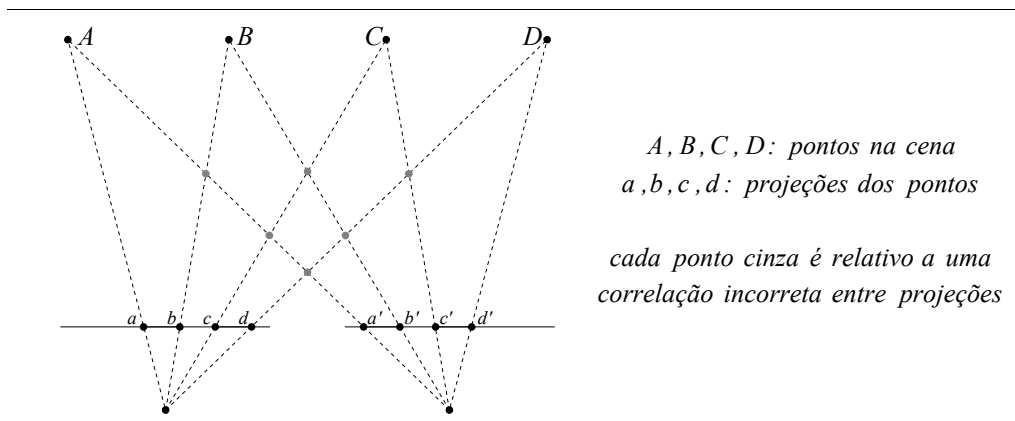


Figura 5. Ambigüidade na correspondência estéreo.

As restrições de unicidade e continuidade foram originalmente propostas por Marr e Poggio [26, 27]. Segundo a restrição de unicidade, cada ponto corresponde a não mais que um ponto na outra imagem, o que não é verdade no caso especial de transparência ou semi-occlusão, onde dois pontos visíveis em uma imagem estão no mesmo raio de projeção na outra.

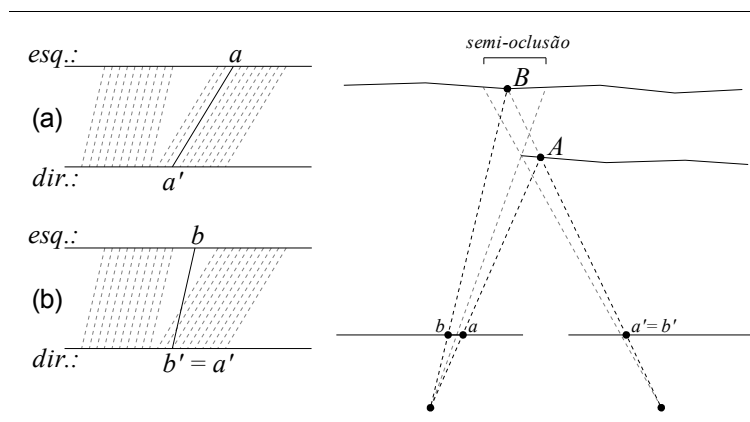


Figura 6. Restrição de unicidade: (a) exemplo; (b) exceção (correlacionamento em áreas de semi-ocusão).

A restrição de continuidade considera que disparidades são geralmente contínuas, uma vez que as profundidades variam suavemente em quase todo o mapa, com a provável exceção das bordas dos objetos, onde podem ocorrer variações abruptas de profundidade. [25]

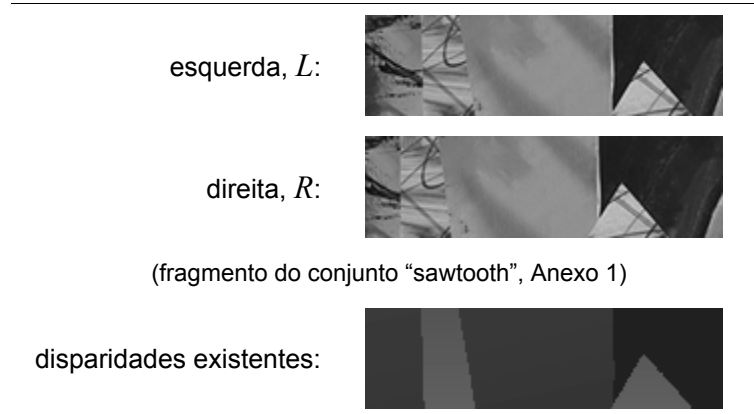


Figura 7. Restrição de continuidade: exemplo (superfície do mesmo objeto); exceção (bordas entre objetos a distâncias diferentes).

A restrição de similaridade, proposta por Grimson [14], supõe que pontos correspondentes possuem valores de intensidade ou vizinhanças semelhantes (fotometria), ou são aspectos com atributos semelhantes (compatibilidade), o que não é sempre verdade para superfícies não-lambertianas², uma vez que as diferenças na forma com que a luz é difundida em diferentes direções pode causar grandes variações nas intensidades luminosas registradas em ambas as câmeras para a mesma superfície observada.

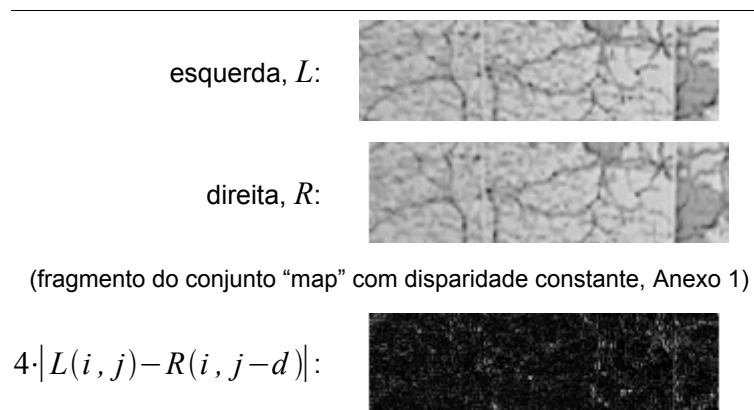


Figura 8. Restrição de similaridade: semelhança fotométrica entre pontos correspondentes (áreas escuras); exceção (pontos mais claros).

² Uma superfície de Lambert é aquela que difunde a luz isotropicamente, isto é, a luz é difundida com intensidade igual em todas as direções.

Baker e Binford propuseram a restrição de ordenação [2], segundo a qual projeções ocorrem na mesma ordem em linhas epipolares correspondentes, o que pode não acontecer quando a distância dos objetos em relação às câmeras variam muito.

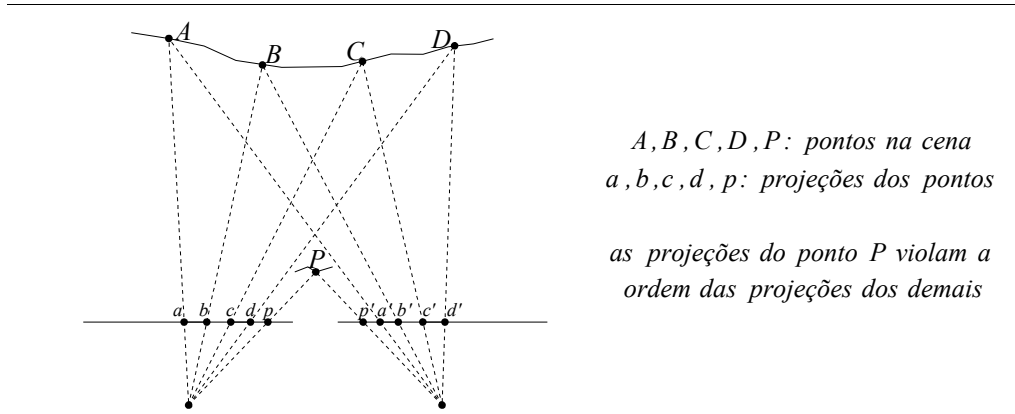


Figura 9. Restrição de ordenação: exemplo (pontos A, B, C, D); exceção (ponto P , distância relativa muito grande entre os objetos).

A restrição de ordenação possui a vantagem de tornar monótona a função de disparidade ao longo de cada linha epipolar. [30]

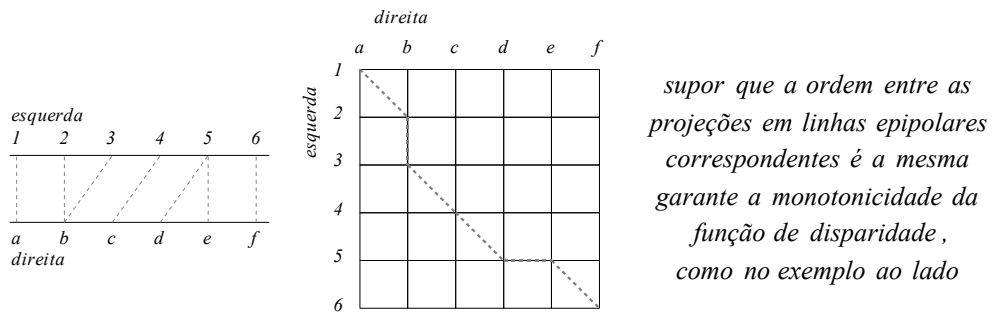


Figura 10. Restrição de ordenação: função de disparidade monótona.

As limitações impostas por essas restrições tornam o problema da correspondência mais específico, simplificando o processo de cálculo das disparidades, uma vez que permitem reduzir a ambigüidade. No entanto, a restrição epipolar, ao contrário dessas restrições fundamentais, é sempre válida para o estabelecimento de correspondências entre um par de imagens estéreo, uma vez que ela é uma restrição geométrica que reduz o espaço de correlação sem excluir possíveis correspondências válidas.

2.3 Trabalhos Relacionados

Métodos para solução do problema da correspondência estéreo podem ser divididos entre aqueles que estabelecem correlações apenas com base em informação local e aqueles que buscam o melhor conjunto de disparidades através de um processo de otimização global.

2.3.1 Métodos Globais

Métodos que calculam disparidades através de um processo de otimização global, via minimização de energia, são de natureza iterativa e, por isso, possuem um custo computacional bastante elevado se comparados a métodos locais. Eles diferem quanto à técnica de minimização utilizada, onde o objetivo é encontrar uma função de disparidade com a menor energia possível. Melhores resultados têm sido obtidos com o uso de propagação de crença [13] e *graph cuts* [20]. A medida de energia avalia o quanto a função de disparidade está de acordo com o par estéreo e o grau de continuidade que ela apresenta. Os métodos propostos por Klaus et al. [19], Q. Yang et al. [43], J. Sun et al. [38] e Bleyer e Gelautz [5] são os mais acurados até o momento, com base na metodologia de teste e conjuntos de dados padrão propostos por Scharstein e Szeliski [35].

Embora possam alcançar resultados de acurácia bastante elevada, métodos que realizam otimização global não são apropriados para uso em sistemas de tempo-real, por causa do processo iterativo que lhes é inerente. Como alternativa ao custo computacional elevado, Q. Yang et al. [42] propõem o uso de *hardware* gráfico (GPU) para reduzir o tempo de processamento (45 vezes, segundo seus resultados) em comparação com a execução não-paralela.

2.3.2 Métodos Locais

Métodos que utilizam apenas informação local objetivam encontrar pares de pontos com a maior similaridade possível entre os dois conjuntos de candidatos de cada imagem. Informação local pode ser considerada na forma de janelas [16], de atributos de aspectos [40] ou de uma união de ambos [8]. Janelas são valores de intensidade na vizinhança e aspectos são padrões bem definidos extraídos previamente da imagem, como retas, arcos, cantos, círculos, elipses etc.

Correlações baseadas em intensidades podem produzir mapas de disparidades bastante densos, mas tendem a gerar erros em áreas onde a textura é repetitiva ou com pouca informação e na bordas dos objetos. Abordagens com janelas adaptativas (de tamanhos e formatos variáveis) foram propostas na tentativa de reduzir esse efeito [18, 41]. Correlações baseadas em aspectos, no entanto, são mais acuradas nas bordas dos objetos, porém geram mapas com disparidades irregularmente distribuídas, além de necessitarem de uma fase de pré-processamento para extração de aspectos.

Métodos locais existentes realizam uma busca para encontrar a melhor correlação para um ponto entre os n pontos na linha epipolar correspondente ou, pelo menos, um subconjunto dela. Isso leva a um processo de natureza combinatória, resultando em uma complexidade não-linear em relação à resolução do par estéreo, uma vez que há um espaço de busca para cada ponto. Uma função de correlação é utilizada para medir a similaridade entre dois pontos com base na informação local. [35]

São exemplos de funções de correlação entre os pontos (i, j) , na imagem esquerda L , e $(i, j - d)$, na imagem direita R , com disparidade d e janela de dimensões M e N :

a) Soma das diferenças absolutas

$$SAD(i, j, d) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |L(i+u, j+v) - R(i+u, j+v-d)|$$

b) Soma das diferenças ao quadrado

$$SSD(i, j, d) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} (L(i+u, j+v) - R(i+u, j+v-d))^2$$

c) Correlação cruzada normalizada

$$NCC(i, j, d) = \frac{\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} (L(i+u, j+v) - \mu(L_{ij})) \cdot (R(i+u, j+v-d) - \mu(R_{ij}))}{\left(\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} (L(i+u, j+v) - \mu(L_{ij}))^2 \cdot \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} (R(i+u, j+v-d) - \mu(R_{ij}))^2 \right)^{\frac{1}{2}}}$$

Onde a função $\mu(I_{ij})$ calcula a média aritmética das intensidades da janela de dimensões M e N com canto superior esquerdo no ponto (i, j) da imagem I :

$$\mu(I_{ij}) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} I(i+u, j+v)$$

A correlação cruzada normalizada não é sensível a mudanças de amplitude da imagem, isto é, mudanças na intensidade média entre as áreas comparadas [22]. Por isso, ela é mais eficiente do que a simples soma das diferenças absolutas ou ao quadrado, embora envolva um número maior de operações.

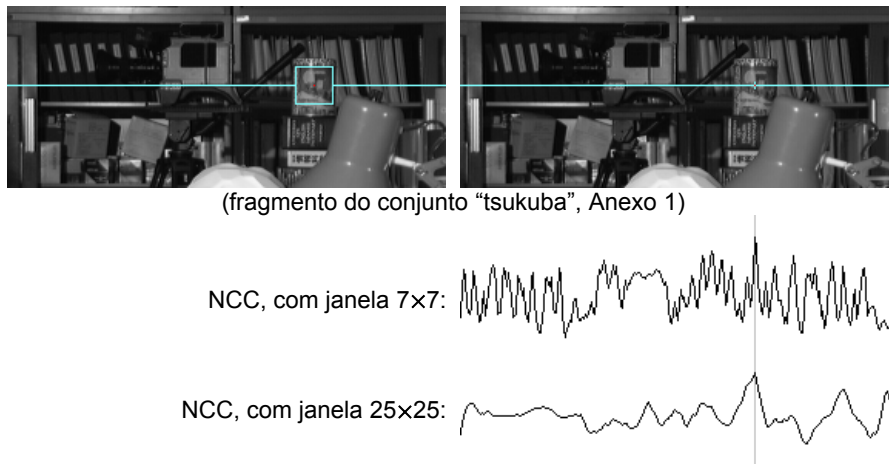


Figura 11. Medida de similaridade para um único ponto ao longo da linha epipolar correspondente (correlação x disparidade).

2.3.3 Relação com Métodos Atuais

Com o objetivo de produzir um mapa de disparidades com acurácia razoável e baixo tempo de processamento, visando à aplicação em sistemas de tempo-real, métodos recentes recorrem a implementação paralela em *hardware* [17], otimização da busca pela melhor correlação [11] ou tentam reduzir o espaço de busca através de esquemas *coarse-to-fine* (correlacionamento em múltiplas resoluções, de imagens pequenas até o tamanho original, iniciando cada passo com disparidades estimadas no passo anterior sobre uma resolução menor) [37] e de predição de janelas de disparidade em uma sequência de imagens estéreo (uma vez que não é esperada grande variação entre um quadro e outro) [28].

Ainda assim, os métodos atuais têm em comum o fato de que realizam uma busca através de um espaço de correlação para cada ponto, a fim de estabelecer uma correspondência na outra imagem. Para que cada disparidade seja calculada em tempo constante, essa busca tem de ser evitada. O cálculo da disparidade de cada ponto em tempo constante permite que o método tenha complexidade linear.

No capítulo 4, onde são discutidos os resultados obtidos, o tempo de processamento do método apresentado é comparado com os tempos de alguns métodos atuais que empregam diferentes técnicas de correlacionamento estéreo. Essa análise é feita para vários conjuntos de dados padrão, estabelecendo claramente a relação linear entre o tempo de processamento e o tamanho das imagens no método apresentado, a qual não é encontrada nas abordagens atuais.

3 Definição do Método

Em métodos locais existentes, uma função de correlação é utilizada para comparar explicitamente duas regiões, resultando em um indicativo do quanto elas são similares e, conseqüentemente, do seu grau de correlação, seja essa informação local tomada como valores de intensidade ou atributos de aspectos. Assim, é feita uma busca pela melhor correlação para cada ponto, com base na informação da região ao seu redor.

No entanto, para evitar a complexidade combinatória inerente a esse processo, a busca pela melhor correlação deve ser evitada. Se não houver uma busca, não há sentido em comparar duas regiões explicitamente, uma vez que cada região poderia ser comparada a apenas uma região na outra imagem.

Uma busca pela melhor correlação possui complexidade $O(n^2)$, se for para cada possível disparidade ao longo da linha epipolar correspondente, ou $O(ns)$, se o tamanho do espaço de busca for restrito a s . Quando $s = 1$, isto é, o espaço é restrito a apenas uma comparação, a complexidade se torna $O(n)$, mas não há como apontar qual é a melhor correlação por meio de uma busca, porque ela não é mais possível. Assim, se não for definido um critério para selecionar uma região dentre as candidatas na outra imagem, é esperada, em média, apenas uma correlação válida em cada linha epipolar. Esse critério tem de se basear no fato de que cada região não pode ser avaliada várias vezes.

3.1 Indexação de Regiões

Um processo de indexação pode tornar possível o estabelecimento de um critério para selecionar uma região correlata com base em como ela é avaliada para o cálculo de seu índice. Assumindo que a função $f(I_{ij})$ calcula um valor discriminante para a região ao redor do ponto (i, j) na imagem I e que f mapeia regiões correlatas para o mesmo valor e não correlatas para valores diferentes, o correlacionamento pode ser feito indexando cada região em ambas as imagens, tomando f como uma função de indexação.

De modo a fazer com que as disparidades ocorram sempre no mesmo sentido, ou seja, tenham o mesmo sinal, o algoritmo utiliza a suposição de que os planos de proje-

ção são coplanares, ou seja, as câmeras são paralelas. Essa suposição permite a simplificação do processo de indexação e serve de base para exclusão de algumas falsas correlações causadas por textura repetitiva ou com pouca informação.

Assim, uma região L_{ij} na imagem esquerda pode ser correlacionada apenas com uma região R_{ij-d} na imagem direita, com uma disparidade $d \geq 0$. Isso permite que todas as candidatas a correlação sejam processadas antes que L_{ij} seja alcançada, permitindo que o processo seja feito em tempo constante para cada ponto ao longo da linha.

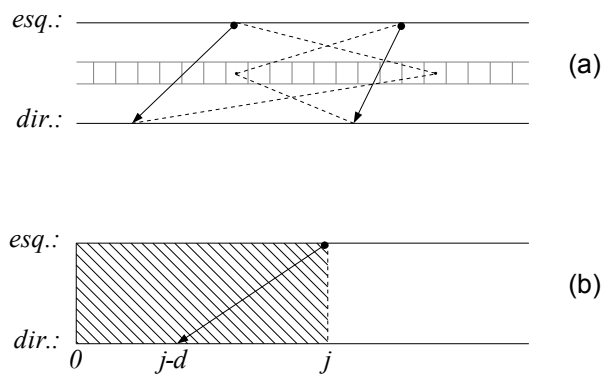


Figura 12. Indexação de regiões: (a) duas correlações; (b) limitação da correlação (disparidade positiva).

A função de indexação atua como uma função heurística, permitindo que um bom número de correlações válidas seja estabelecido em um algoritmo de menor complexidade, comparado àqueles que realizam uma busca pela melhor correlação. O tempo de processamento é independente do espaço de disparidade e da informação contida na imagem, enquanto que a quantidade de correlações válidas é dependente da função de indexação.

Assumindo que o par estéreo está retificado, ou seja, que as linhas epipolares coincidem com as linhas horizontais em ambas as imagens, o processo de indexação ocorre para cada linha, percorrendo-a do início ao fim e tomando um par de regiões, uma de cada imagem, a cada passo. Uma região na imagem direita é indexada se a posição no vetor de indexação estiver livre e uma região na imagem esquerda é relacionada com uma região previamente indexada se a posição estiver ocupada por ela.

```

para linha  $i$  de 0 até  $m-1$ 
   $p \leftarrow 0$ 
  para coluna  $j$  de 0 até  $n-1$ 
    se (vetor( $f(R_{ij})$ ) está livre)
      vetor( $f(R_{ij})$ )  $\leftarrow j$ 
      lista( $p$ )  $\leftarrow f(R_{ij})$ 
       $p \leftarrow p+1$ 
    fim
    se (vetor( $f(L_{ij})$ ) não está livre)
      disparidade( $i, j$ )  $\leftarrow j - \text{vetor}(f(L_{ij}))$ 
      libera vetor( $f(L_{ij})$ )
    fim
  fim
para posição  $j$  de 0 até  $p-1$ 
  se (vetor(list( $j$ )) não está livre)
    libera vetor(lista( $j$ ))
  fim
fim

```

Figura 13. Algoritmo básico para indexação de regiões.

Uma região R_{ij-d} , na imagem direita, pode ser incorretamente correlacionada com a região $L_{ij-d'}$, na imagem esquerda, antes da região correspondente L_{ij-d} (que seria a correlação válida) ser alcançada, onde $0 < d' \leq d$, se a área de $j-d$ a j for de textura repetitiva ou com pouca informação, uma vez que os valores discriminantes tendem a ter pouca variação nessas circunstâncias. Por isso, é comum que d' seja próximo ou igual a d .

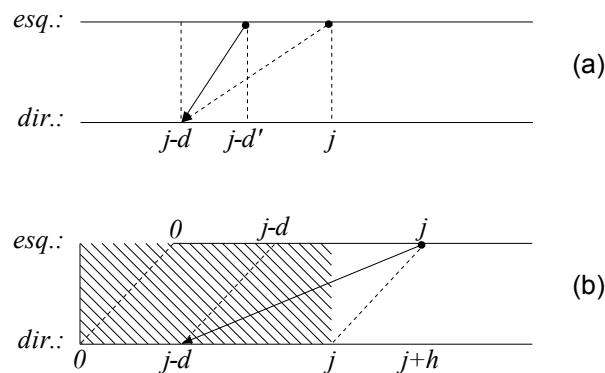


Figura 14. Problema de repetição do discriminante:
(a) falsa correlação; (b) deslocamento horizontal.

Pode ser difícil detectar essas falsas correlações, mas muitas delas podem ser evitadas ainda no processo de indexação, através do igual deslocamento de todas as regiões na imagem esquerda, de modo que L_{ij} possa ser correlacionada com regiões até a coluna $j+h$, onde h é um valor de deslocamento positivo. Dessa forma, correlações que resultarem em disparidades negativas devem ter se originado por causa da pouca variação de textura e podem ser ignoradas, uma vez que as disparidades sempre ocorrem no mesmo sentido, conforme a restrição colocada na Figura 12b.

```

para coluna  $j$  de  $-h$  até  $n-1$ 
  se  $(j+h < n) \wedge (\text{vetor}(f(R_{ij+h})) \text{ está livre})$ 
     $\text{vetor}(f(R_{ij+h})) \leftarrow j+h$ 
     $\text{lista}(p) \leftarrow f(R_{ij+h})$ 
     $p \leftarrow p+1$ 
  fim
  se  $(j \geq 0) \wedge (\text{vetor}(f(L_{ij})) \text{ não está livre})$ 
    se  $(j - \text{vetor}(f(L_{ij})) \geq 0)$ 
       $\text{disparidade}(i, j) \leftarrow j - \text{vetor}(f(L_{ij}))$ 
    fim
    libera  $\text{vetor}(f(L_{ij}))$ 
  fim
fim

```

Figura 15. Modificação do algoritmo para indexação em uma linha deslocada.

Enquanto uma região da imagem direita está indexada e esperando ser correlacionada, sua posição no vetor de indexação permanece ocupada. Assim, outras regiões que eventualmente possuam o mesmo discriminante não são indexadas e perdem a oportunidade de serem correlacionadas.

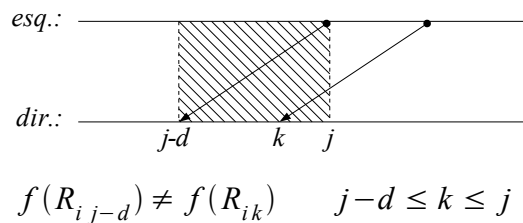


Figura 16. Indisponibilidade do índice.

3.1.1 Função de Indexação

A função de indexação deve mapear regiões similares para o mesmo valor e não similares para valores diferentes, onde ser similar significa possuir um padrão em comum. Também, disparidades resultantes de falsas correlações devem compor um conjunto de maior entropia, comparado a disparidades resultantes de correlações válidas, para que seja mais fácil fazer uma separação posteriormente.

Supondo que $nc(L, R)$ seja uma função de correlação normalizada para um valor no intervalo $[0, 1]$ que estima o grau de similaridade entre as regiões L e R , as três características acima podem ser definidas como o seguinte:

(i) se duas regiões L_{ij} e R_{ij-d} são similares, é grande a probabilidade de que seus valores discriminantes sejam os mesmos, ou seja, similaridade implica discriminantes iguais na maioria dos casos:

$$\begin{aligned} nc(L_{ij}, R_{ij-d}) &\approx 1 \\ \rightarrow P(f(L_{ij}) = f(R_{ij-d})) &\approx 1 \end{aligned} \quad (1)$$

(ii) se duas regiões L_{ij} e $R_{ij-\tilde{d}}$ não são similares, é grande a probabilidade de que seus valores discriminantes sejam diferentes, ou seja, não similaridade implica discriminantes diferentes na maioria dos casos:

$$\begin{aligned} nc(L_{ij}, R_{ij-\tilde{d}}) &\approx 0 \\ \rightarrow P(f(L_{ij}) \neq f(R_{ij-\tilde{d}})) &\approx 1 \end{aligned} \quad (2)$$

(iii) se duas regiões L_{ij} e $R_{ij-\tilde{d}}$ não são similares e seu valor discriminante é o mesmo, então essa falsa correlação resulta em um valor de disparidade $\tilde{d} \in [0, j]$ seguindo uma distribuição de probabilidade uniforme:

$$\begin{aligned} nc(L_{ij}, R_{ij-\tilde{d}}) &\approx 0 \quad \wedge \quad f(L_{ij}) = f(R_{ij-\tilde{d}}) \\ \rightarrow P(\tilde{d} = x) &\approx \frac{1}{j+1} \quad 0 \leq x \leq j \end{aligned} \quad (3)$$

A característica (iii) fornece a base para uma detecção robusta de falsas correlações através de restrição de continuidade. Ela coloca que, se uma correlação não for válida, a disparidade resultante pode assumir qualquer valor possível com aproximadamente a mesma probabilidade, uma vez que qualquer mudança no valor discriminante pode levar a uma correlação totalmente arbitrária.

3.1.2 Indexação Baseada em Intensidades

A função de indexação $f(I_{ij})$, utilizada para produzir os resultados experimentais aqui apresentados, é baseada em apenas alguns dos valores de intensidade da região I_{ij} para que seja possível calcular o valor discriminante com um número pequeno de operações. O cálculo é feito com base na média aritmética dos valores da região e de como os pontos selecionados diferem dela.

O termo principal da função f é o índice de região y_{reg} . Cada bit nesse índice indica se o valor de intensidade correspondente está acima ou abaixo da média da região:

$$y_{reg}(I_{ij}) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \left[2^{\sigma_K(u,v)-1} \cdot K(u,v) \cdot \delta(I(i+u, j+v), \mu(I_{ij})) \right] \quad (4)$$

O *kernel* K define quais pontos da região são utilizados no cálculo do índice, onde $K(u, v) \in \{0, 1\}$, $0 \leq u < M$, $0 \leq v < N$. A função $\sigma_K(u, v)$ informa o número de pontos iguais a 1 no *kernel* K , linha por linha, até a posição (u, v) :

$$\sigma_K(u, v) = \sum_{w=0}^{u \cdot N + v} K(w \setminus N, r) \quad (5)$$

$$w \equiv r \pmod{N}$$

A função de limiar $\delta(x, t)$ tem saídas 1 se $x \geq t$ e 0 para qualquer outro caso:

$$\delta(x, t) = \begin{cases} 1 & x \geq t \\ 0 & x < t \end{cases} \quad (6)$$

A função $\mu(I_{ij})$ calcula a intensidade média da região I_{ij} , a qual é usada como valor de limiar em (4):

$$\mu(I_{ij}) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} I(i+u, j+v) \quad (7)$$

Dessa forma, um bit no índice de região é colocado como 1 apenas se o valor de intensidade correspondente for maior ou igual à média da região. Apenas pontos iguais a 1 no *kernel* são utilizados nesse cálculo. O número de bits no índice é igual ao número de pontos na região. A Figura 17 apresenta um exemplo de como ele é calculado a partir de uma região 4×4.

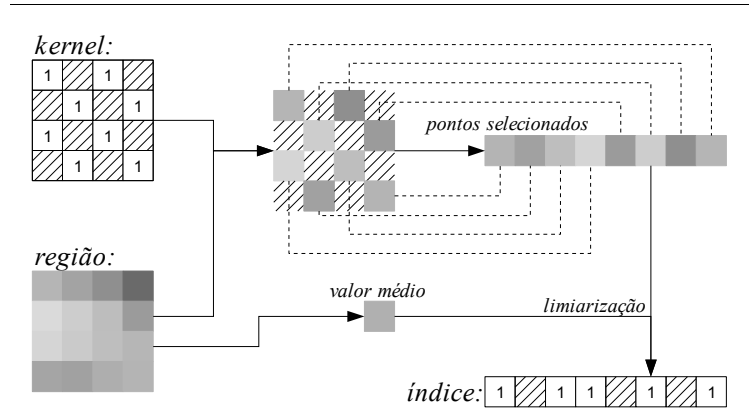


Figura 17. Exemplo de cálculo do índice de região.

Aumentar ou diminuir igualmente todos os valores de intensidade não afeta o valor do índice de região. Regiões com pouca similaridade podem ter o mesmo índice em consequência disso, mas elas podem ser separadas utilizando sua intensidade média. Se o vetor de indexação for segmentado com base nesse valor, apenas regiões com o mesmo índice e uma intensidade média aproximada poderão ser correlacionadas. Se forem desejados 2^S segmentos, S será o tamanho, em bits, do índice de segmento y_{seg} .

A segmentação do vetor de indexação com base na intensidade média da região ajuda a espalhar as indexações ao longo do vetor, o que permite reduzir as colisões e aumentar potencialmente o número de correlações.

A função de indexação f concatena os bits desses dois índices, gerando assim o valor discriminante da região, com o índice de segmento ocupando os bits mais significativos:

$$f(I_{ij}) = y_{seg}(I_{ij}) \cdot 2^{\sigma_K(M-1, N-1)} + y_{reg}(I_{ij}) \quad (8)$$

Onde $\sigma_K(M-1, N-1)$ informa o tamanho, em bits, do índice de região y_{reg} , ou seja, o número de pontos iguais a 1 no *kernel* K , os quais serão utilizados para o cálculo do índice, e $y_{seg}(I_{ij})$ é o índice de segmento de I_{ij} :

$$y_{seg}(I_{ij}) = \mu(I_{ij}) \setminus 2^{D-S} \quad (9)$$

Onde D é o tamanho, em bits, de um valor de intensidade, ou seja, a profundidade dos tons de cinza, de modo que 2^D seja o número de possíveis intensidades. O índice de segmento y_{seg} equivale aos S bits mais significativos da intensidade média da região, ou seja, é feito um deslocamento dos bits para a direita.

Tabela 1. Parâmetros do algoritmo de indexação de regiões.

deslocamento lateral, h	número de colunas a serem deslocadas, conforme descrito na Figura 14b
tamanho do índice de segmento, S	número de bits do índice que endereça um segmento no vetor de indexação – consequentemente, define o número de segmentos
área de uma região, $M \times N$	tamanho fixo da área retangular que compreende uma região
pontos utilizados, K	matriz binária que define quais pontos da região entram no cálculo do índice – tem o mesmo tamanho da área de uma região

Essa função de indexação é baseada apenas em intensidades e usa uma janela de tamanho fixo, da qual nem todos os pontos precisam ser avaliados para calcular do valor discriminante utilizado no processo de correlacionamento.

3.2 Detecção de Falsas Correlações

No processo de indexação, falsas correlações resultam em disparidades com maior entropia do que aquelas resultantes de correlações verdadeiras, uma vez que, no caso da correlação ser falsa, a disparidade pode assumir qualquer valor possível com a mesma probabilidade, de acordo com a característica (iii) da função de indexação, descrita formalmente em (3). Ao ser aplicada uma restrição de continuidade ao mapa de disparidades, é possível detectar falsas correlações em um processo de regularização com base na variação de entropia.

A aprovação de cada ponto no mapa de disparidades depende do ponto e de sua vizinhança estarem de acordo com a restrição de continuidade definida em (10). Cada disparidade d possui um peso w_d , que é a média aritmética das frequências de disparidades afins em todo o mapa, ou seja, qualquer disparidade s , tal que $|s - d| \leq 1$. Assim, o vetor de pesos W é uma suavização do histograma do mapa. Uma correlação é considerada válida se o somatório dos pesos de disparidades afins estiver acima de um percentual mínimo do somatório dos pesos de todas as disparidades na vizinhança:

$$\sum_{s=d-1}^{d+1} v_s \cdot w_s \geq \left(\sum_{a=0}^{n-1} v_a \cdot w_a \right) \cdot (1 - \tau) \quad (10)$$

Onde τ é um valor de tolerância de descontinuidade no intervalo $[0, 1]$, V é o histograma da janela de verificação (vizinhança) e n é o número de colunas da imagem (disparidades variam de 0 a $n-1$). Ainda, como critério adicional, uma correlação pode ser removida se o número de disparidades que são iguais a d estiver abaixo de um certo valor mínimo q .

A restrição de continuidade é aplicada para cada ponto no mapa. Se não houver uma disparidade calculada para um certo ponto, é utilizada a última disparidade avaliada. Assim, áreas esparsas podem ser reduzidas a partir da suposição de que uma disparidade próxima ainda pode ser aprovada para a nova vizinhança, como se resultasse de uma correlação válida.

O mapa resultante pode ser equalizado durante o processo de regularização se cada disparidade aprovada for substituída pela média ponderada das disparidades afins na vizinhança:

$$\bar{d} = \frac{\sum_{s=d-1}^{d+1} v_s \cdot w_s \cdot s}{\sum_{s=d-1}^{d+1} v_s \cdot w_s} \quad (11)$$

A equalização torna o mapa mais uniforme em áreas onde a diferença entre as disparidades é pequena.

Tabela 2. Parâmetros do algoritmo de detecção de falsas correlações.

janela de verificação, $L_v \times L_h$	tamanho fixo da área retangular que compreende a vizinhança de um ponto para teste de continuidade
tolerância de descontinuidade, τ	percentual máximo de descontinuidade permitido – valor entre 0 e 1
mínimo de disparidades iguais, q	número mínimo de disparidades na vizinhança que devem ser iguais ao ponto sendo testado

As Figuras de 18 a 21 descrevem o algoritmo de aplicação da restrição de continuidade para um mapa de m linhas e n colunas, onde L_v e L_h são constantes e correspondem, respectivamente, às dimensões vertical e horizontal da janela de verificação.

Inicialmente, é calculado o histograma ponderado U para a janela relativa ao canto superior esquerdo do mapa, tal que $U(x) = V(x) \cdot W(x)$, onde V é o histograma das disparidades na janela e W é o vetor de pesos, ou seja, o histograma suavizado de todo o mapa tomado como entrada.

A janela de verificação é deslocada para a direita até o final da linha e, então, é deslocada para a linha de baixo para que possa percorrê-la até seu início. Assim, o mapa é processado em ziguezague, sendo necessário apenas deslocar a janela em uma coluna ou linha para avaliar cada novo ponto. Em cada deslocamento, o vetor U é atualizado de acordo com os pontos que entram e saem do alcance da janela.

```

calcular o histograma da janela em (0,0)
para linha i de 0 até m-Lv
  se (i é ímpar)
    para coluna j de 0 até n-Lh-1
      aplicar a restrição para o ponto em (i, j)
      atualizar o histograma para a janela em (i, j+1)
    fim
    aplicar a restrição para o ponto em (i, n-Lh)
  se (i < m-Lv)
    atualizar o histograma para a janela em (i+1, n-Lh)
  fim
senão
  para coluna j de n-Lh até 1
    aplicar a restrição para o ponto em (i, j)
    atualizar o histograma para a janela em (i, j-1)
  fim
  aplicar a restrição para o ponto em (i, 0)
  se (i < m-Lv)
    atualizar o histograma para a janela em (i+1, 0)
  fim
fim

```

Figura 18. Algoritmo para aplicação da restrição de continuidade.

A cada deslocamento da janela de verificação, além do histograma ponderado, é atualizado o somatório dos pesos de todos os pontos dentro da janela, que é usado no teste de continuidade em (10) e equivale a $\sum_{a=0}^{n-1} v_a \cdot w_a$ e a $\sum_{a=0}^{n-1} u_a$. Ele é atualizado também de acordo com os pontos que entram e saem do alcance da janela.

<pre> para j de 1 até n-2 W(d) ← H(j-1) + H(j) + H(j+1) fim W(0) ← $\frac{H(0) + H(1)}{2}$ W(n-1) ← $\frac{H(n-2) + H(n-1)}{2}$ </pre>	<pre> soma ← 0 para linha i de 0 até L_v para coluna j de 0 até L_h d ← disparidade(i, j) se (d foi calculada) U(d) ← U(d) + W(d) soma ← soma + W(d) fim fim fim </pre>
--	---

Figura 19. Cálculo do vetor de pesos W a partir da suavização do histograma do mapa de entrada (a esquerda) e cálculo do histograma ponderado U para a primeira janela (a direita).

```

 $d \leftarrow entrada(i, j)$ 
se ( $d$  foi calculada)
     $\bar{d} \leftarrow d$ 
fim
 $similares \leftarrow U(\bar{d}) + U(\bar{d}+1)$ 
se ( $\bar{d} > 0$ )
     $similares \leftarrow similares + U(\bar{d}-1)$ 
fim
se ( $soma > 0$ )  $\wedge$  ( $similares \geq (1-\tau) \cdot soma$ )  $\wedge$  ( $U(\bar{d}) \geq q \cdot W(\bar{d})$ )
    se (deve equalizar)
        se ( $\bar{d} > 0$ )
             $saída(i, j) \leftarrow \frac{\bar{d} \cdot U(\bar{d}) + (\bar{d}+1) \cdot U(\bar{d}+1) + (\bar{d}-1) \cdot U(\bar{d}-1)}{similares}$ 
        senão
             $saída(i, j) \leftarrow \frac{\bar{d} \cdot U(\bar{d}) + (\bar{d}+1) \cdot U(\bar{d}+1)}{similares}$ 
        fim
    senão
         $saída(i, j) \leftarrow \bar{d}$ 
    fim
fim

```

Figura 20. Aplicação da restrição de continuidade para um ponto (i, j) .

O teste de continuidade para cada ponto é feito com base na relação entre o somatório dos pesos das disparidades similares e o somatório dos pesos de todos os pontos dentro da janela. A inequação $U(\bar{d}) \geq q \cdot W(\bar{d})$ equivale a $V(\bar{d}) \geq q$ e é utilizada no lugar da segunda, uma vez que V não é explicitamente conhecido ao longo do processo.

$(i, j) \rightarrow (i, j+1)$ para x de i até $i+L_v-1$ $d \leftarrow entrada(x, j)$ se (d foi calculada) $w \leftarrow W(d)$ $U(d) \leftarrow U(d) - w$ $soma \leftarrow soma - w$ fim $d \leftarrow entrada(x, j+L_h)$ se (d foi calculada) $w \leftarrow W(d)$ $U(d) \leftarrow U(d) + w$ $soma \leftarrow soma + w$ fim fim	$(i, j) \rightarrow (i, j-1)$ para x de i até $i+L_v-1$ $d \leftarrow entrada(x, j+L_h-1)$ se (d foi calculada) $w \leftarrow W(d)$ $U(d) \leftarrow U(d) - w$ $soma \leftarrow soma - w$ fim $d \leftarrow entrada(x, j-1)$ se (d foi calculada) $w \leftarrow W(d)$ $U(d) \leftarrow U(d) + w$ $soma \leftarrow soma + w$ fim fim	$(i, j) \rightarrow (i+1, j)$ para x de j até $j+L_h-1$ $d \leftarrow entrada(i, x)$ se (d foi calculada) $w \leftarrow W(d)$ $U(d) \leftarrow U(d) - w$ $soma \leftarrow soma - w$ fim $d \leftarrow entrada(i+L_v, x)$ se (d foi calculada) $w \leftarrow W(d)$ $U(d) \leftarrow U(d) + w$ $soma \leftarrow soma + w$ fim fim
---	---	---

Figura 21. Atualização do histograma para janelas em $(i, j+1)$, $(i, j-1)$ e $(i+1, j)$, respectivamente.

3.3 Interpolação do Mapa de Disparidades

Após a aplicação da restrição de continuidade, áreas esparsas remanescentes são preenchidas através de um algoritmo de interpolação. Para que esse processo não tome muito tempo de processamento, cada ponto no mapa tem o valor da disparidade mais próxima na mesma linha ou coluna ou adjacentes.

									3	3	3									
									2	2	2									
	8	7	6	5	4	3	2	1	1	1	2	3	4	5	6	7	8			
	8	7	6	5	4	3	2	1		1	2	3	4	5	6	7	8			
	8	7	6	5	4	3	2	1	1	1	2	3	4	5	6	7	8			
									2	2	2									
									3	3	3									

Figura 22. Linhas e colunas usadas na interpolação com base no ponto mais próximo.

Essa interpolação baseada no ponto mais próximo é limitada ortogonalmente, mas pode gerar resultados razoavelmente uniformes, mesmo em casos onde o ponto mais próximo está consideravelmente distante, por causa da boa continuidade do mapa.

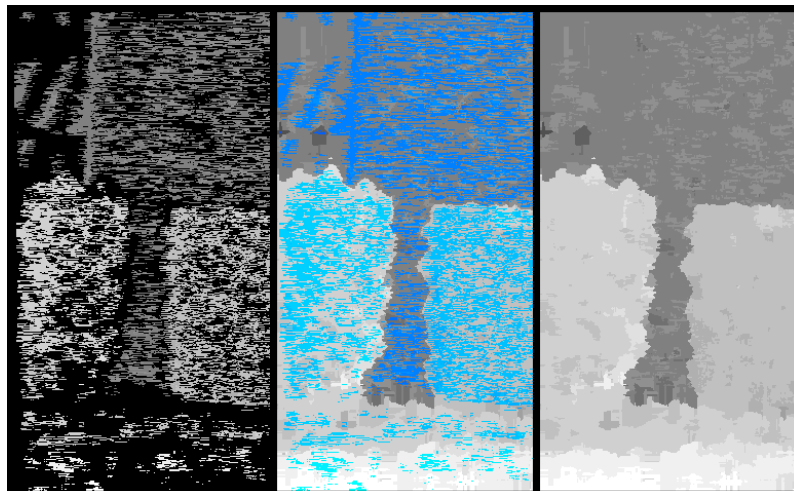


Figura 23. Exemplo de interpolação com base no ponto mais próximo (disparidades originais à esquerda e em azul).

<pre> para j de 0 até n-1 abaixo(j).disparidade ← valor indefinido abaixo(j).distância ← ∞ fim para j de 0 até n-1 acima(j).disparidade ← valor indefinido acima(j).distância ← ∞ fim </pre>	<pre> d ← entrada(i, j) se (d não está definida) d ← entrada(i-1, j) se (d não está definida) d ← entrada(i+1, j) se (d não está definida) d ← entrada(i, j-1) se (d não está definida) d ← entrada(i, j+1) fim fim fim fim fim </pre>
---	--

Figura 24. Algoritmo de interpolação: inicialização (à esquerda); obtenção da disparidade para um ponto (i, j) (à direita).

```

para linha i de m-1 até 0
  disparidade ← valor indefinido
  distância ← ∞
  para coluna j de n-1 até 0
    obter disparidade para o ponto (i, j)
    se (d está definida)
      disparidade ← d
      distância ← 0
      abaixo(j).disparidade ← d
      abaixo(j).distância ← 0
    senão
      se (distância < ∞)
        distância ← distância + 1
      fim
      se (abaixo(j).distância < ∞)
        abaixo(j).distância ← abaixo(j).distância + 1
      fim
    fim
  se (distância ≤ abaixo(j).distância)
    temp(i, j).disparidade ← disparidade
    temp(i, j).distância ← distância
  senão
    temp(i, j).disparidade ← abaixo(j).disparidade
    temp(i, j).distância ← abaixo(j).distância
  fim
fim
fim

```

Figura 25. Algoritmo de interpolação: geração do mapa intermediário.

O mapa de disparidades é processado para gerar um mapa intermediário, onde cada ponto contém a distância e o respectivo valor da disparidade mais próxima nas

duas direções ortogonais (ou seja, é considerada a última disparidade visitada na mesma linha ou adjacente e a última disparidade visitada na mesma coluna ou adjacente).

```

para linha  $i$  de 0 até  $m-1$ 
  disparidade  $\leftarrow$  valor indefinido
  distância  $\leftarrow \infty$ 
  para coluna  $j$  de 0 até  $n-1$ 
    obter disparidade para o ponto  $(i, j)$ 
    se ( $d$  está definida)
      disparidade  $\leftarrow d$ 
      distância  $\leftarrow 0$ 
      abaixo( $j$ ).disparidade  $\leftarrow d$ 
      abaixo( $j$ ).distância  $\leftarrow 0$ 
    senão
      se (distância  $< \infty$ )
        distância  $\leftarrow$  distância + 1
      fim
      se (acima( $j$ ).distância  $< \infty$ )
        acima( $j$ ).distância  $\leftarrow$  acima( $j$ ).distância + 1
      fim
    fim
  se (distância  $\leq$  acima( $j$ ).distância)
    se (distância  $\leq$  temp( $i, j$ ).distância)
      saída( $i, j$ )  $\leftarrow$  disparidade
    senão
      saída( $i, j$ )  $\leftarrow$  temp( $i, j$ ).disparidade
    fim
  senão
    se (acima( $j$ ).distância  $\leq$  temp( $i, j$ ).distância)
      saída( $i, j$ )  $\leftarrow$  acima( $j$ ).disparidade
    senão
      saída( $i, j$ )  $\leftarrow$  temp( $i, j$ ).disparidade
    fim
  fim
fim

```

Figura 26. Algoritmo de interpolação: geração do mapa final.

O mapa final é calculado percorrendo o mapa original no sentido contrário, mantendo apenas a disparidade com a menor distância nas quatro direções.

4 Resultados

Nos resultados a partir de imagens reais, aproximadamente metade das regiões na imagem direita foram indexadas. Isso ocorre por causa de colisões no vetor de indexação, uma vez que regiões que possuam o mesmo discriminante não podem ser indexadas ao mesmo tempo, havendo um período de indisponibilidade do índice, conforme exibido na Figura 16. Ainda, aproximadamente metade das regiões indexadas foram correlacionadas com regiões na imagem esquerda. Resultados intermediários, usando o conjunto de dados “shrub” [7], são exibidos nas Figuras de 27 a 30.

Um filtro de média aritmética com janela 2×2 é aplicado antes da indexação de regiões, porque um maior número de regiões são indexadas e correlacionadas quando o par de imagens é levemente suavizado. Intensidades individuais podem variar entre regiões correspondentes por causa de diferenças na maneira com que a mesma superfície é projetada em cada uma das câmeras. Essa variação pode afetar o índice de região se o valor da intensidade estiver próximo da média da região. Suavizar o par de imagens faz com que cada bit do índice de região seja baseado em mais de um ponto da imagem original, reduzindo o impacto dessas variações.

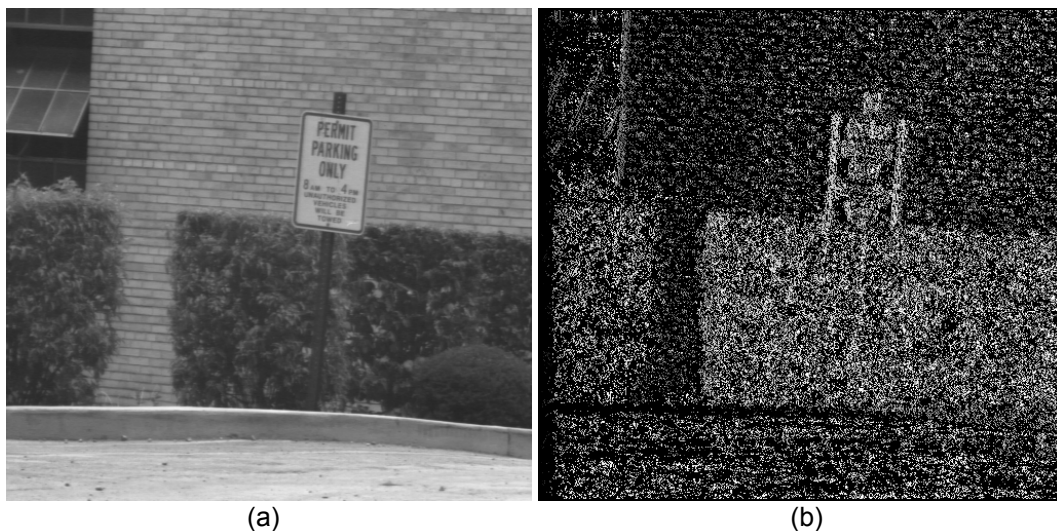


Figura 27. Resultado da indexação de regiões: (a) imagem esquerda; (b) disparidades calculadas (indexações: 51%, correlações: 27%);

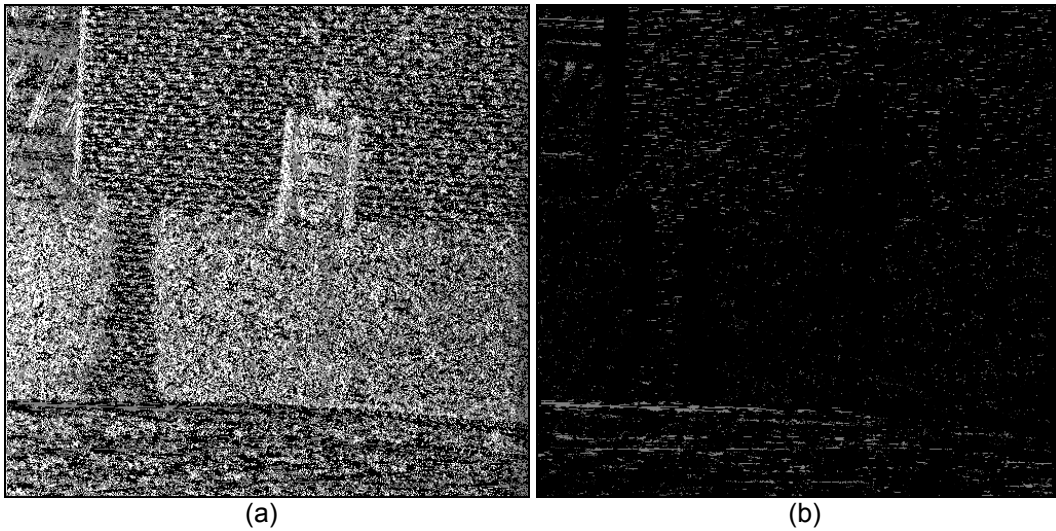


Figura 28. Mapas de eficiência da indexação de regiões:
(a) colisões (em cinza) e correlações (em branco);
(b) disparidades negativas (rejeitadas durante a indexação);

O mapa calculado no processo de indexação de regiões é semi-denso e possui uma grande quantidade de falsas correlações, como pode ser observado na Figura 27b. Após a aplicação da restrição de continuidade, o mapa resultante apresenta disparidades bem distribuídas e com entropia menor, assim como uma quantidade consideravelmente menor de falsas correlações, como exibido na Figura 30.

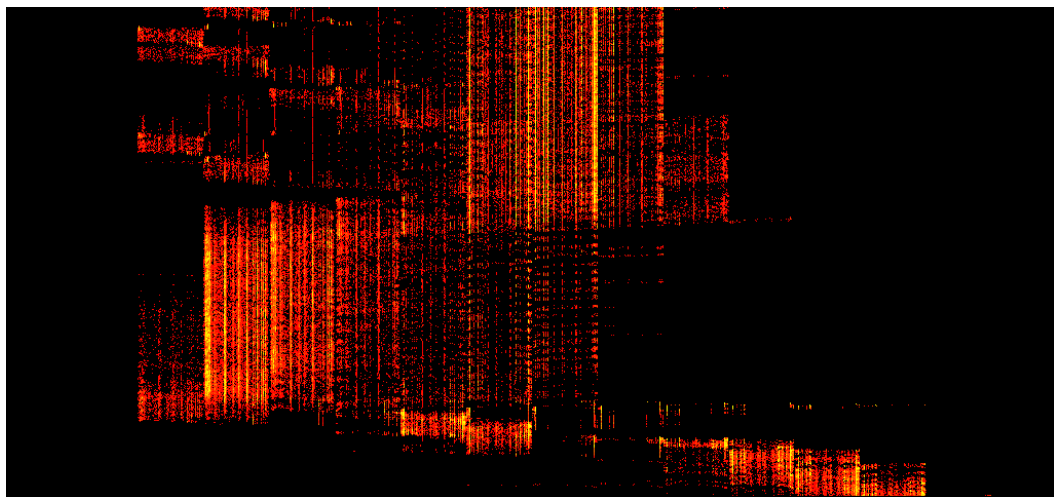


Figura 29. Histograma de ocupação dos vetores de indexação para cada linha epipolar (cada linha da imagem corresponde a um vetor).

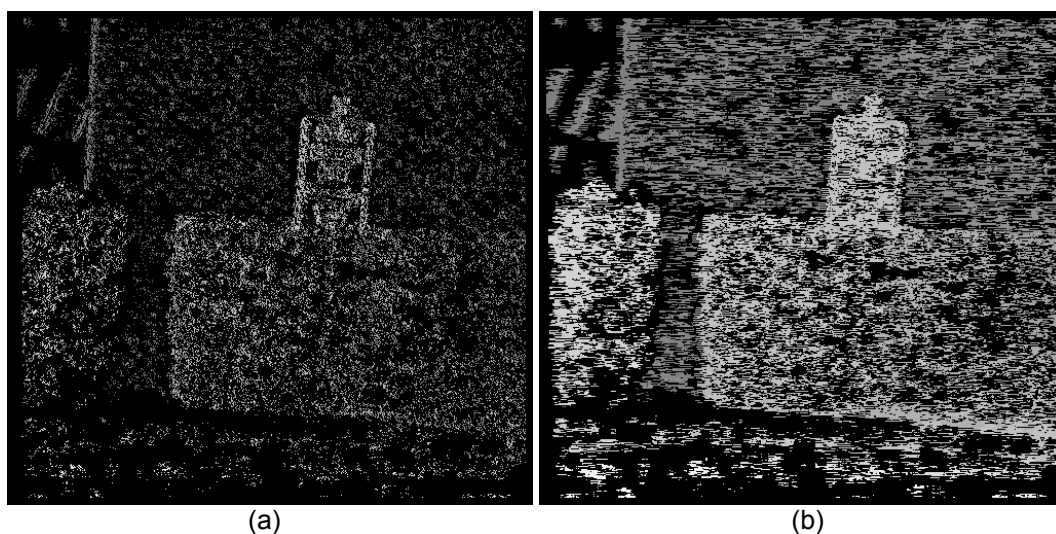


Figura 30. Resultado após a restrição de continuidade:
(a) para as disparidades calculadas na indexação (densidade: 13%);
(b) para todos os pontos no mapa (densidade: 40%).

O objetivo da restrição de continuidade é remover falsas correlações (redução de falsos positivos) e aumentar a densidade do mapa (redução de falsos negativos). Um mapa semi-denso, mas com disparidades bem distribuídas, ainda pode ter sua densidade aumentada através de interpolação. No entanto, as falsas correlações são um problema maior, cuja solução depende, em grande parte, da aplicação dessa restrição.

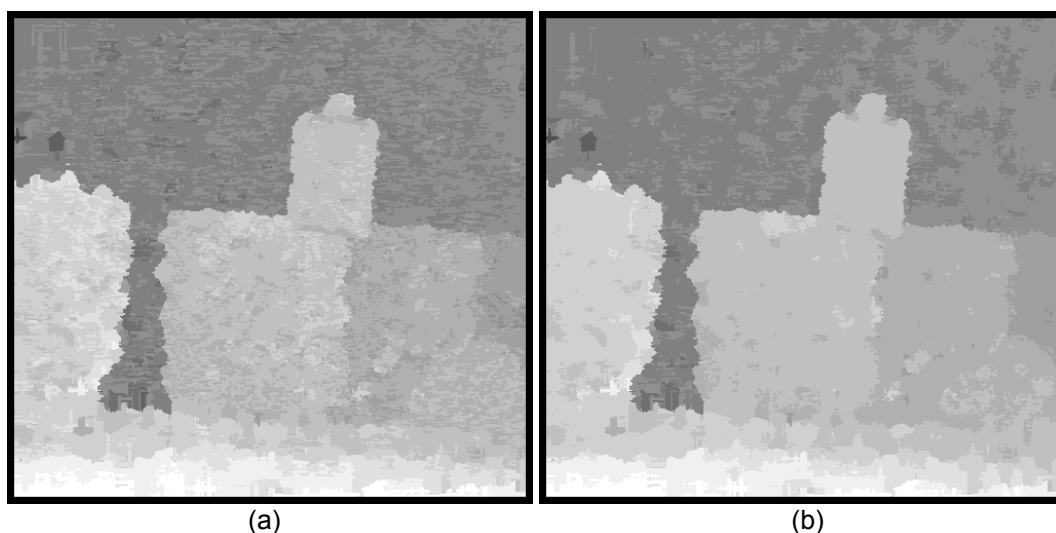


Figura 31. Interpolação baseada no ponto mais próximo:
(a) mapa não equalizado; (b) mapa equalizado.

O método produz mais de 30 quadros por segundo com uma resolução de 256×256 em um computador pessoal comum³ (menos de um microsegundo por ponto). Isso é feito para qualquer informação de disparidade contida no par estéreo, enquanto que nos métodos atuais isso não ocorre, uma vez que essa informação afeta o tamanho do espaço de busca e, conseqüentemente, o tempo de processamento.

Foi observado empiricamente, a partir dos conjuntos de dados apresentados mais adiante, que o tempo de processamento do método depende apenas da resolução do par de imagens, como proposto em teoria.

4.1 Parametrização

Todos os resultados apresentados neste trabalho foram produzidos com a mesma parametrização, isto é, não foi feito qualquer ajuste nos valores dos parâmetros para casos particulares. Isso demonstra que o método pode tomar diferentes conjuntos de dados de entrada e ainda produzir resultados satisfatórios sem a intervenção do usuário, o que é imprescindível para sua aplicação em sistemas de tempo-real.

Tabela 3. Valores padrão para os parâmetros do método.

deslocamento lateral, h	8 colunas
tamanho do índice de segmento, S	4 bits
área de uma região, $M \times N$	4×4
pontos utilizados, K	$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$
janela de verificação, $L_v \times L_h$	15×15
tolerância de descontinuidade, τ	0,6
mínimo de disparidades iguais, q	8

³ A configuração do computador pessoal utilizado está detalhada no Anexo 2.

O deslocamento lateral h torna maior o período de indisponibilidade do índice, o que aumenta a chance de ocorrer uma falsa correlação, mas permite que falsas correlações causadas pela pouca variação de textura possam ser ignoradas ainda na fase de indexação de regiões, uma vez que ele tem o objetivo de produzir disparidades negativas em áreas onde ocorre repetição do valor discriminante. Por isso, um pequeno número de colunas deslocadas já é suficiente.

A segmentação do vetor de indexação é baseada na suposição de que regiões correspondentes possuem uma intensidade média igual ou semelhante (de acordo com a restrição de similaridade). Um número muito grande de segmentos não é interessante porque deve ser permitida uma certa variação da intensidade média da região sem que o índice de segmento seja alterado, porém um número muito pequeno reduz a eficiência da segmentação. Por isso, um tamanho S razoável para o índice de segmento é a metade do número de bits do valor de uma intensidade luminosa, o que equivale a 4 (uma variação de 16) para intensidades representadas por inteiros de 8 bits.

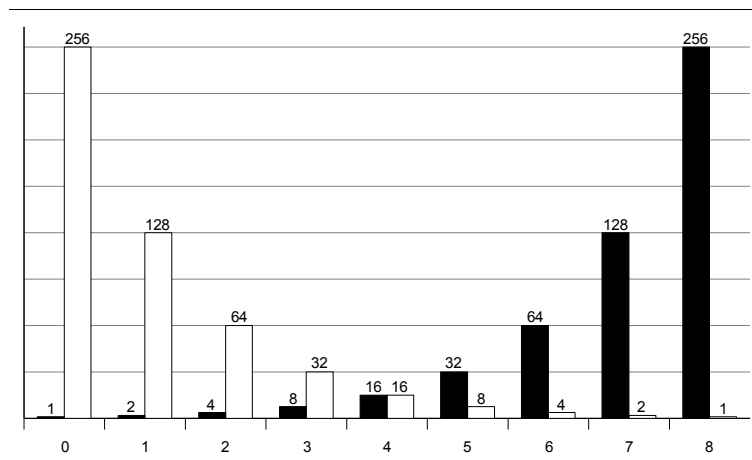


Figura 32. Número de segmentos (em preto) e variação (em branco) em relação ao tamanho S do índice de segmento (para intensidades de 8 bits).

O número de pontos utilizados no cálculo do índice de região não pode ser muito pequeno para que seja obtida informação suficiente da região. No entanto, quanto maior for o número de pontos tomados, maior é a chance de que uma diferença considerável em um desses pontos cause uma mudança no valor discriminante em relação à região correspondente.

4.2 Metodologia de Teste

O teste de acurácia dos resultados foi feito seguindo a metodologia de avaliação proposta por Scharstein e Szeliski [35], na qual o erro é calculado através da subtração entre o mapa de disparidades calculadas pelo método e o mapa de disparidades esperadas. Assim, o valor do erro é o percentual de todas as diferenças absolutas maiores do que um certo limiar de erro (geralmente de 1 ponto), sem contar os pontos em áreas de oclusão. O mapa com a definição de tais áreas e o mapa de disparidades esperadas são conhecidos e fazem parte de cada conjunto de dados, além do próprio par estéreo.

4.3 Conjuntos de Dados Padrão

O método foi testado com os seguintes conjuntos de dados padrão contendo mapa de disparidades esperadas e mapa de oclusão:

- “tsukuba”, “sawtooth”, “venus” e “map”; [35]
- “cones” e “teddy”. [36]

O conjunto “tsukuba” foi introduzido por Nakamura et al. [29]; o conjunto “map” por Szeliski e Zabih [39]; e os demais por Scharstein e Szeliski [35, 36].

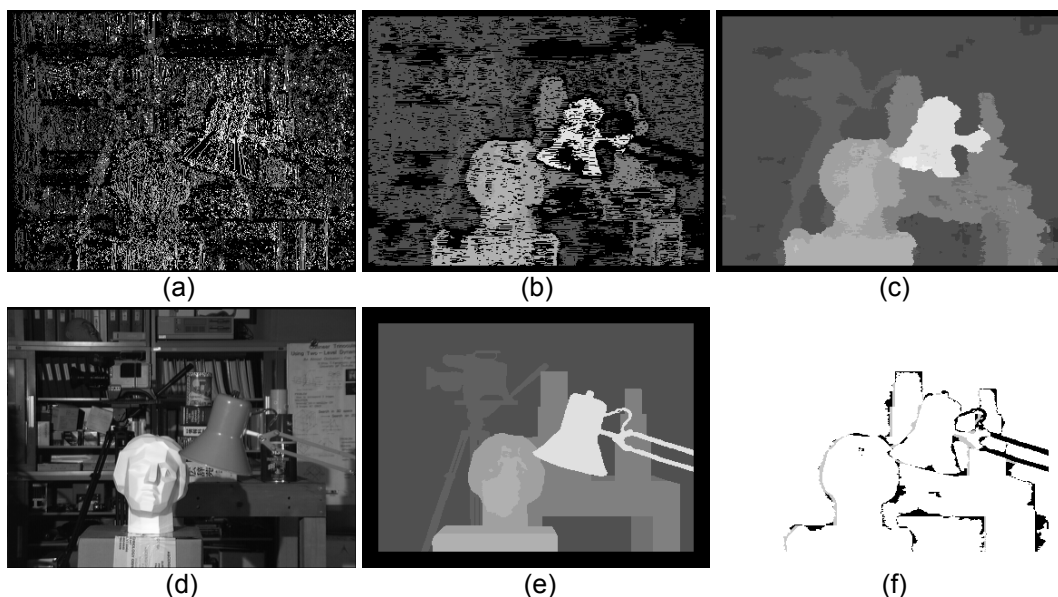


Figura 33. Resultados no conjunto “tsukuba”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).

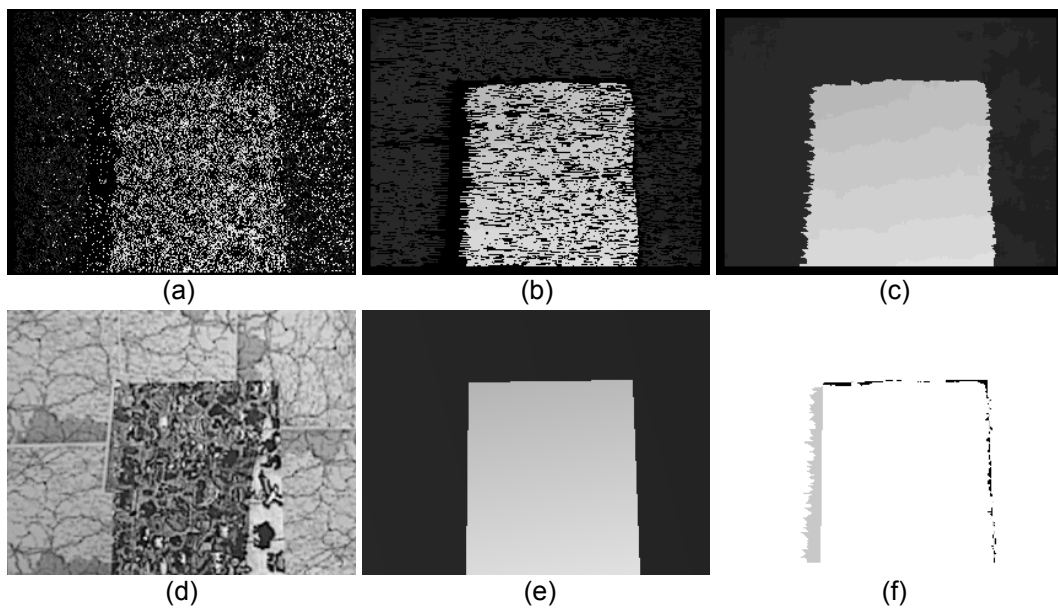


Figura 34. Resultados no conjunto “map”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).

Grande parte do erro se concentra na borda dos objetos, o que faz com que objetos muito finos se percam, como pode ser visto na Figura 33c (a haste dupla da luminária não foi preservada). A maior parte do erro na Figura 34f ocorre por causa da existência de pontos semi-occludidos (em cinza), uma vez que o método, embora remova corretamente as falsas correlações em áreas de oclusão, não detecta tais pontos durante o processo, o que impede uma interpolação mais acurada.

A quantidade elevada de disparidades incorretas calculadas pelo algoritmo de indexação de regiões é visível na Figura 33a. A grande entropia do conjunto de disparidades resultantes de falsas correlações permite uma eficiente remoção dessas disparidades através da restrição de continuidade. Essa diferença de entropia é potencialmente maior do que em métodos que usam uma função de correlação direta entre duas regiões para estabelecer a correspondência através de uma busca, uma vez que falsas correlações podem resultar em valores de disparidade próximos dos corretos e não com uma distribuição de probabilidade uniforme como ocorre neste método.

A baixa densidade do mapa de disparidades após a indexação de regiões evidencia uma sensibilidade não muito alta do método, mas a eficiente remoção de falsas correlações permite uma especificidade considerável.

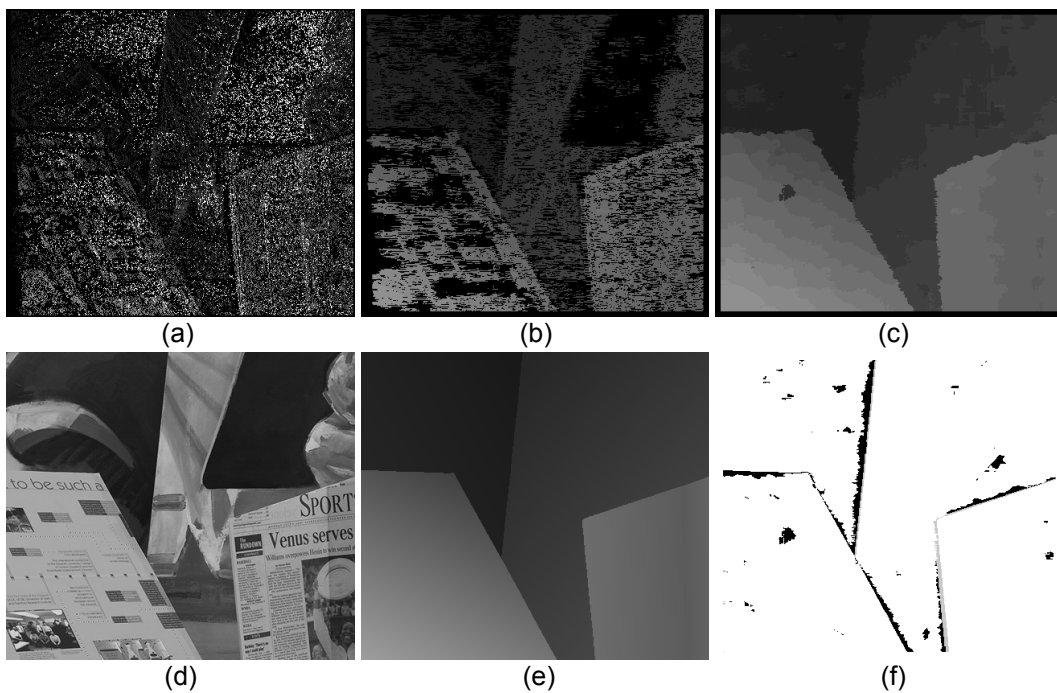


Figura 35. Resultados no conjunto “venus”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).

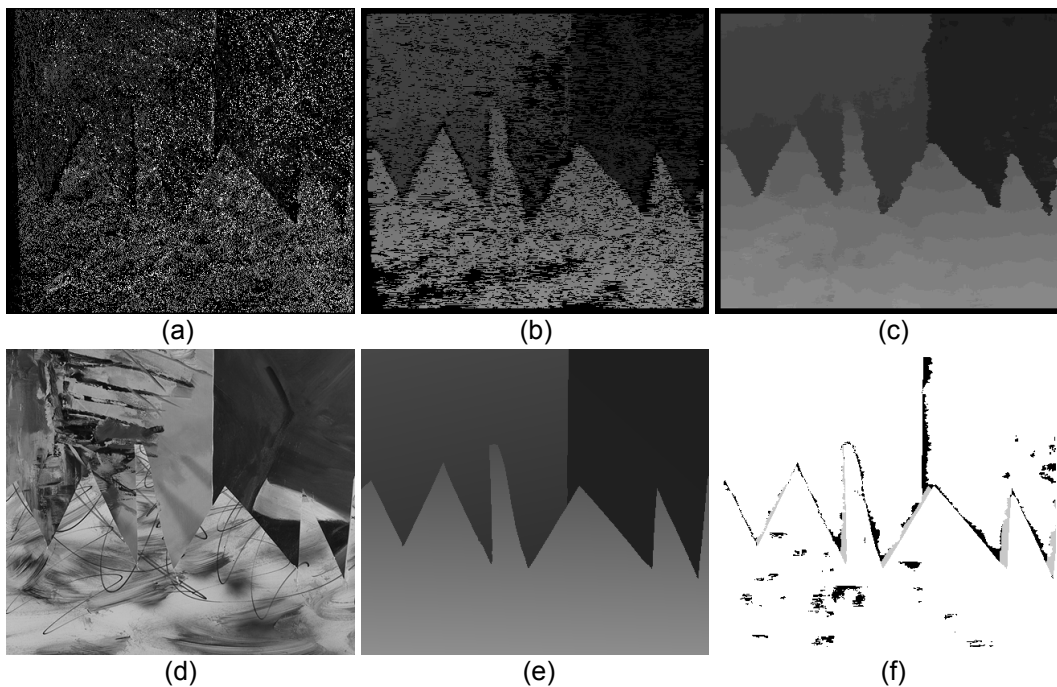


Figura 36. Resultados no conjunto “sawtooth”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).

Alguns pontos de erro na Figura 35f são causados por textura repetitiva. Como o índice de região é baseado em uma suposição local de similaridade (isto é, regiões com o mesmo valor discriminante são provavelmente correlatas), repetições de textura geram densos grupos de falsas correlações, os quais podem ser de difícil remoção através da restrição de continuidade, embora sua maior parte possa ser ignorada na fase de indexação de regiões se for utilizada a modificação do algoritmo exibida na Figura 15.

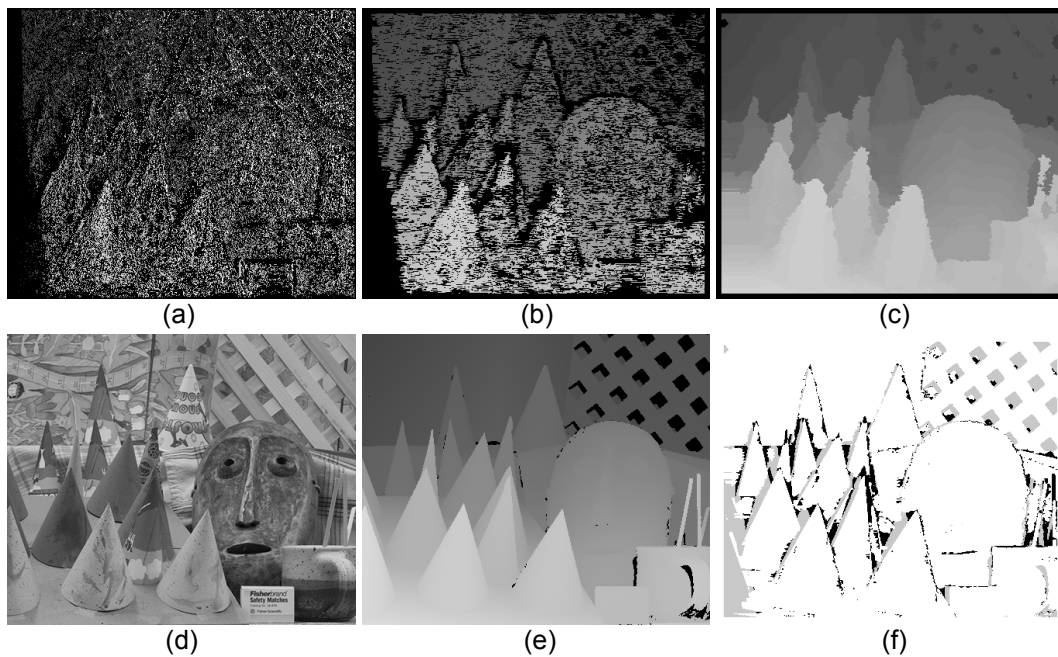


Figura 37. Resultados no conjunto “cones”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).

Os resultados exibidos nas Figuras 37 e 38 são os que apresentam maior erro, principalmente por causa da maior complexidade da cena observada. Um parte considerável do erro encontrado na Figura 38f está situado na parte inferior do mapa, que é uma área relativa ao chão da cena e, por conta da inclinação e da pouca textura, concentrou muitas falsas correlações e acabou ficando bastante esparsa após a restrição de continuidade.

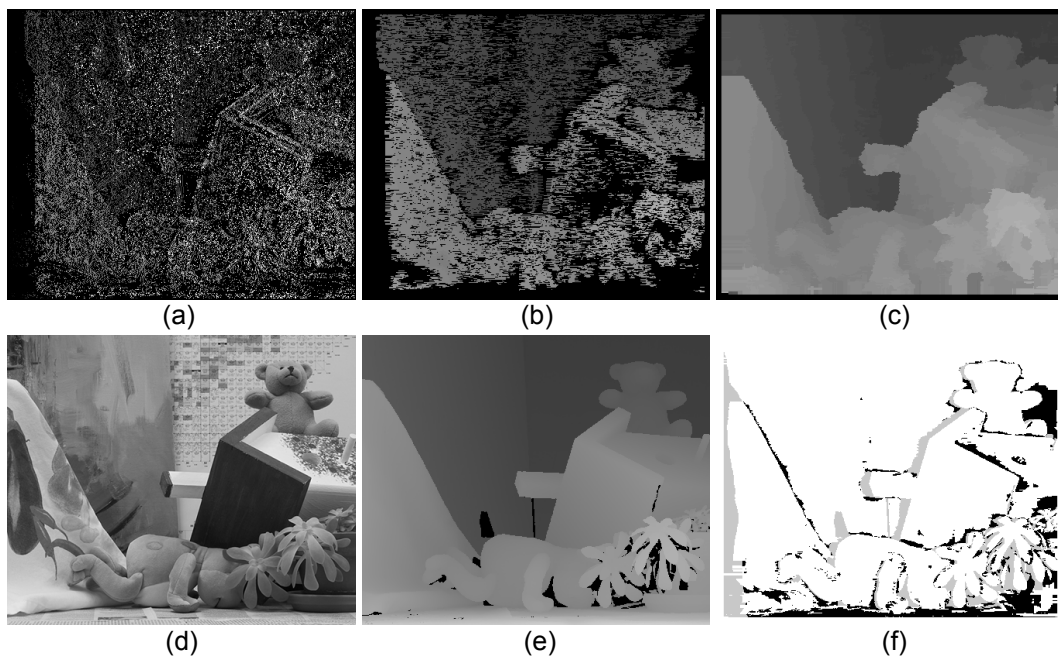


Figura 38. Resultados no conjunto “teddy”: (a) indexação de regiões; (b) restrição de continuidade; (c) interpolação; (d) imagem esquerda; (e) disparidades esperadas; (f) mapa de erro (oclusão em cinza).

A tabela abaixo exhibe dados sobre as fases intermediárias do método, quando aplicado a cada conjunto de dados padrão, incluindo o percentual de erro para pontos sem oclusão e o tempo de processamento de cada fase.

Tabela 4. Acurácia e tempo de processamento.

	<i>map</i>	<i>tsukuba</i>	<i>sawtooth</i>	<i>venus</i>	<i>cones</i>	<i>teddy</i>
resolução	284×216	384×288	434×380	434×383	450×375	450×375
indexação						
regiões	59853	108585	162487	163780	166284	166284
indexadas	79%	67%	72%	72%	71%	68%
correlacionadas	35%	35%	37%	34%	35%	33%
tempo	10 ms	17 ms	27 ms	27 ms	27 ms	27 ms
continuidade						
disp. válidas	25%	25%	27%	22%	22%	20%
densidade final	66%	59%	64%	55%	54%	51%
tempo	15 ms	26 ms	42 ms	41 ms	42 ms	41 ms
interpolação						
tempo	2 ms	4 ms	6 ms	6 ms	6 ms	6 ms
erro	0,63%	4,07%	3,33%	3,23%	5,68%	9,91%
tempo total	27 ms	47 ms	75 ms	74 ms	75 ms	74 ms

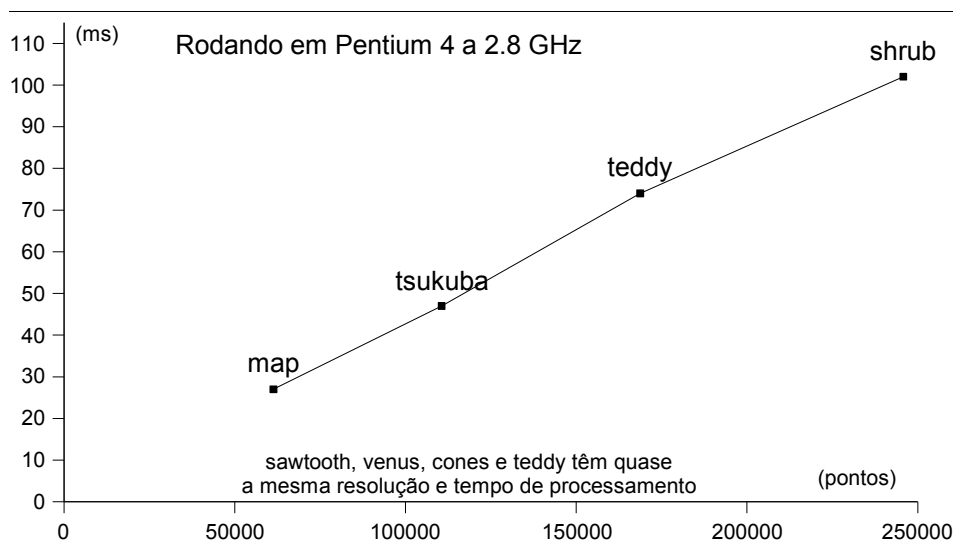


Figura 39. Tempo de processamento para conjuntos de dados com diferentes resoluções.

Como cada ponto é processado em tempo constante, o tempo de processamento depende apenas da resolução do par de imagens. Essa proporção linear pode ser observada nos resultados exibidos na Figura 39.

4.3.1 Comparativo entre Métodos

A Figura 40 exibe um gráfico com os tempos de processamento do método apresentado e de outros frequentemente utilizados em sistemas de tempo-real. As implementações utilizadas para a realização desses testes são uma parte do módulo de avaliação de métodos de visão estéreo desenvolvido por Scharstein e Szeliski [35].

As abordagens comparadas são:

- RI – indexação de regiões;
- SAD – soma das diferenças absolutas;
- SSD – soma das diferenças ao quadrado;
- DP – programação dinâmica;
- SO – programação dinâmica assimétrica.

Na abordagem DP, cada linha é otimizada por meio de programação dinâmica com aplicação da restrição de ordenação, ao passo que na SO (*scanline optimization*), a otimização se dá de maneira assimétrica (sem restrição de ordenação) e mediante uma medida de suavidade horizontal. Sem essa medida, o processo seria equivalente a tomar

sempre a melhor correlação (maior similaridade), isto é, realizar uma busca simples pela melhor correlação. Para este último caso, são comparados métodos utilizam funções de correlação SAD e SSD.

Os quatro métodos utilizados na comparação foram testados tendo a informação prévia do tamanho máximo do espaço de disparidade em cada conjunto de dados. Por isso, os tempos de processamento obtidos por eles são os melhores possíveis, uma vez que a busca foi feita apenas no espaço entre a mínima e a máxima disparidade contida no par de imagens. O método baseado em indexação de regiões não possui essa limitação de ser dependente do espaço de disparidade e apresentou um tempo de processamento bastante inferior aos demais.

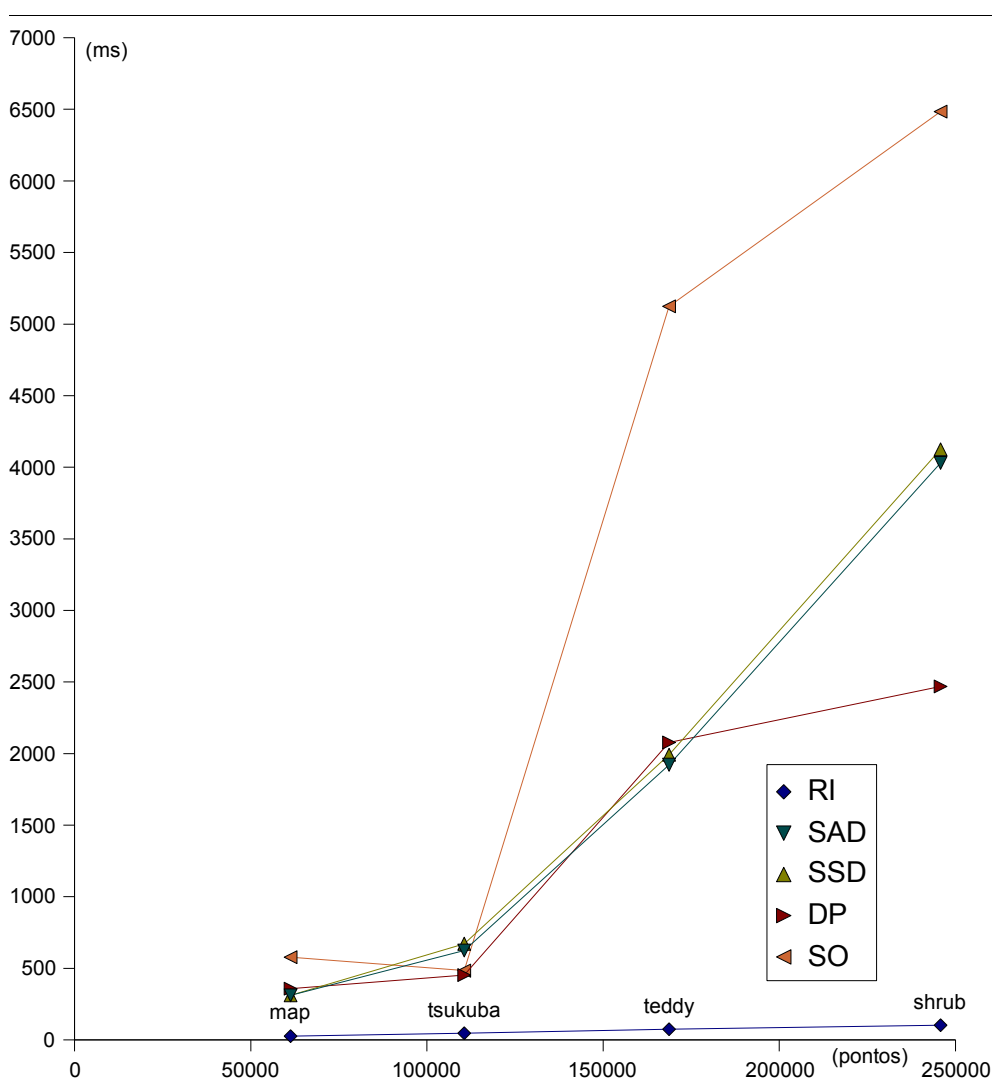


Figura 40. Tempos de processamento de diferentes métodos para conjuntos de dados com diferentes resoluções.

5 Conclusões

O método demonstrou boa robustez por não necessitar de ajuste de parâmetros durante os testes com uma gama variada de conjuntos de dados de entrada, com acurácia mínima de 90% em todos os testes. Sua complexidade linear permite que o tempo de processamento não seja dependente do espaço de disparidade, uma vez que não é feita uma busca pela melhor correspondência para cada ponto. Também não há limites para o espaço de disparidade, ou seja, é possível calcular disparidades de qualquer tamanho até a largura da imagem. Essa é geralmente uma limitação comum em outros métodos, onde um espaço máximo é arbitrado previamente para reduzir o tempo de processamento dessa busca.

Para que uma região possa ser correlacionada, todas as candidatas na linha epipolar correspondente devem ser indexadas primeiro. Por isso, todas as disparidades têm de ocorrer no mesmo sentido, o que faz com que o algoritmo de correlacionamento estéreo seja limitado a uma entrada obtida a partir de planos de projeção coplanares (câmeras paralelas). Portanto, é preciso assegurar que, ao retificar o par estéreo, as imagens não apresentem disparidades nos dois sentidos, pois as negativas serão suprimidas durante o processo de indexação.

O mapa resultante é semi-denso, mas as disparidades são bem distribuídas e as áreas esparsas, causadas por oclusão ou pouca textura, são preenchidas satisfatoriamente através de interpolação baseada no ponto mais próximo. Embora o método apresente uma sensibilidade baixa, comparado a outros métodos locais baseados em correlação de área, sua especificidade é grande, removendo grande parte das falsas correlações resultantes do algoritmo de indexação.

Como consequência da falta de tratamento específico de oclusão e também da simplicidade da função de indexação, as bordas dos objetos não são bem definidas e concentram a maior parte do erro quando em comparação com o resultado esperado.

Além da complexidade linear, a simplicidade da função de indexação e do processo de correlacionamento por indexação de regiões permitem que o método possa ser fa-

cilmente utilizado em sistemas de tempo-real, demandando menos de um microsegundo para processar cada ponto em um computador pessoal comum.

Ainda que métodos baseados em otimização, por sua natureza iterativa, não sejam indicados para sistemas de tempo-real, sua grande acurácia evidencia que a visão estéreo passiva (usando apenas a informação contida no par estéreo, sem a necessidade de agir sobre o meio através de emissões de padrões luminosos) pode gerar resultados bastante próximos daqueles obtidos através da ativa (usando luz estruturada ou varredura a laser). Isso permite supor que, como os métodos de visão passiva buscam uma solução para o mesmo problema mal-colocado, a acurácia dos atuais métodos locais de menor complexidade, como o apresentado neste trabalho, ainda pode ser aumentada. Uma maior clareza sobre novos caminhos a serem seguidos depende dos resultados de novas abordagens para o problema da correspondência e do refinamento dos métodos locais existentes.

5.1 Trabalhos Futuros

Os trabalhos futuros podem ser separados quanto a refinamentos (aumento da acurácia) e extensão do método para o caso mais geral do problema da correspondência.

5.1.1 Refinamentos

Os possíveis refinamentos são:

(a) O método apresentado utiliza uma função de indexação baseada em intensidades luminosas bastante simples e, por isso, computacionalmente eficiente. Uma função de indexação baseada em atributos de aspectos pode ajudar a reduzir o erro nas bordas dos objetos, porém a extração de aspectos e a própria função devem ser eficientes para que a substituição seja interessante;

(b) Embora a segmentação do vetor de indexação promova um bom espalhamento dos índices no vetor de indexação (como pode ser visto na Figura 29), um melhor espalhamento pode levar a uma redução do número de colisões durante o processo de correlacionamento e, conseqüentemente, a um aumento do número de correlações;

(c) Uma melhoria da detecção de falsas correlações através de uma comparação direta (como SAD, SSD etc.), após a correspondência ter sido estabelecida, pode ser interessante se a eficiência dessa validação compensar seu custo;

(d) A segmentação do mapa de disparidades pode ser utilizada na fase de interpolação para detectar áreas de oclusão e ajudar a reduzir distorções em tais áreas.

5.1.2 Extensão

O método apresentado se propõe a resolver o problema da correspondência estéreo (disparidade em apenas uma direção). Para o caso mais geral do problema da correspondência, a restrição epipolar não é aplicável, uma vez que as disparidades podem ocorrer em qualquer direção bidimensional (de um ponto em uma imagem para qualquer ponto na outra). Assim sendo, uma extensão do algoritmo de correlacionamento por indexação de regiões para esse caso mais geral deve obter como solução um mapa de disparidades bidimensionais. Para tanto, um único vetor de indexação poderia ser utilizado para toda a imagem, o que potencialmente degradaria a acurácia.

Como consequência dessa generalização, o método poderia ser utilizado para detecção de continuidade temporal e estimativa de movimento a partir da correspondência livre entre dois quadros, que é uma aplicação para a qual ainda não há métodos de complexidade linear.

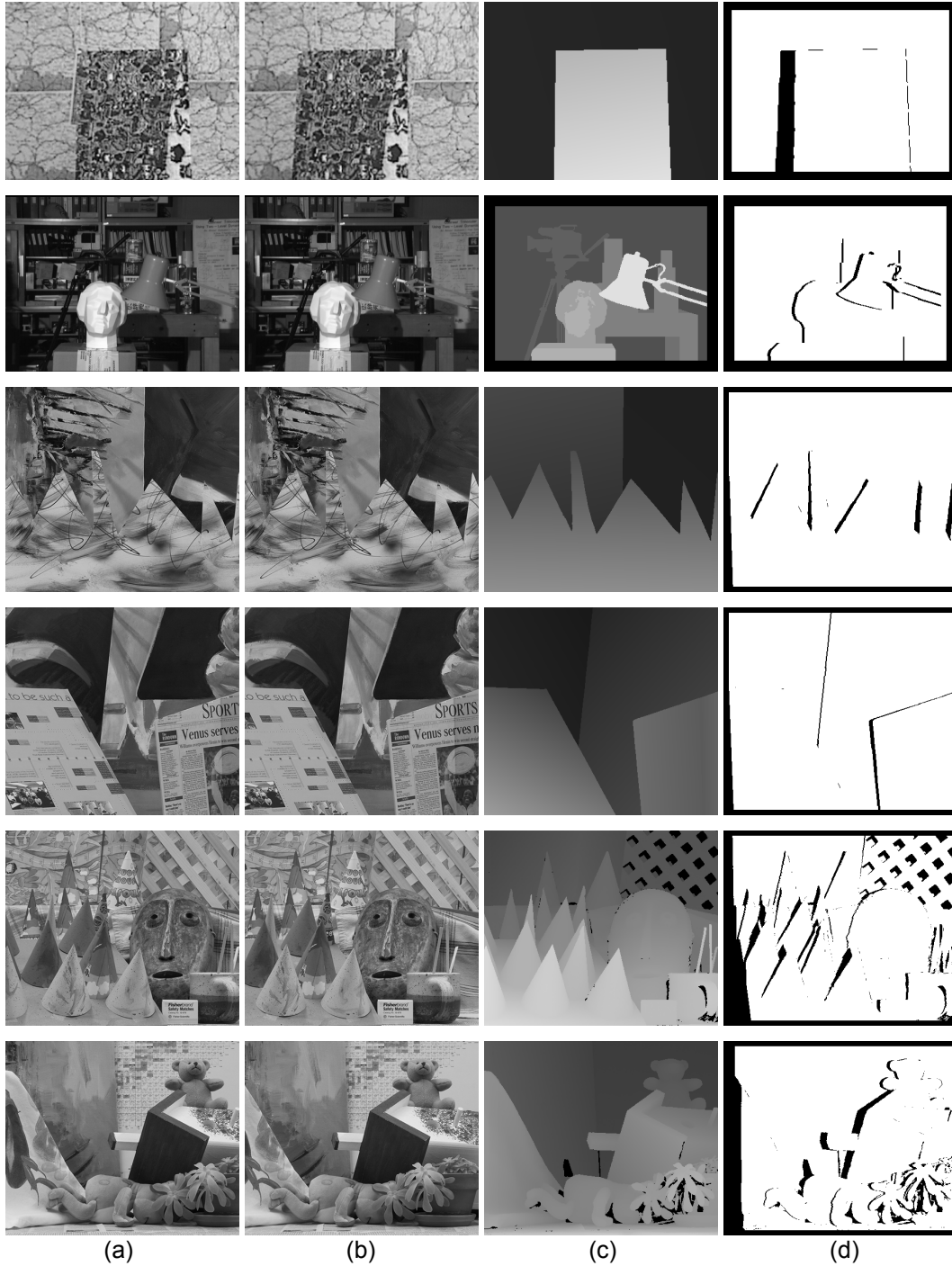
Referências Bibliográficas

- [1] AYACHE, N.; HANSEN, C. **Rectification of Images for Binocular and Trinocular Stereovision**. 9th International Conference on Pattern Recognition, v. 1, p. 11-16, 1988.
- [2] BAKER, H. H.; BINFORD, T. O. **Depth from Edge and Intensity Based Stereo**. 7th International Joint Conference on Artificial Intelligence, p. 631-636, 1981.
- [3] BARNARD, S. T.; FISCHLER, M. A. **Computational Stereo**. ACM Computing Surveys, v. 14, p. 553-572, 1982.
- [4] BLEYER, M.; GELAUTZ, M. **A Layered Stereo Algorithm Using Image Segmentation and Global Visibility Constraints**. IEEE International Conference on Image Processing, v. 5, p. 2997-3000, 2004.
- [5] BLEYER, M.; GELAUTZ, M. **Graph-Based Surface Reconstruction from Stereo Pairs Using Image Segmentation**. International Society for Optical Engineering, v. 5665, pp. 288-299, 2005.
- [6] BRADY, M. J. **Computational Approaches to Image Understanding**. ACM Computing Surveys, v. 14, p. 3-71, 1982.
- [7] BOLLES, R. C.; BAKER, H. H.; HANNAH, M. J. **The JISCT Stereo Evaluation**. DARPA Image Understanding Workshop, p. 263-274, 1993.
- [8] COCHRAN, S. D.; MEDIONI, G. **3-D Surface Description from Binocular Stereo**. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 14, p. 981-994, 1992.
- [9] DAVIS, J.; RAMAMOORTHY, R.; RUSINKIEWICZ, S. **Spacetime Stereo: A Unifying Framework for Depth from Triangulation**. IEEE Conference on Computer Vision and Pattern Recognition, v. 2, p. 359-366, 2003.
- [10] DERICHE, R.; ZHANG, Z.; LUONG Q.-T.; FAUGERAS, O. **Robust Recovery of the Epipolar Geometry for an Uncalibrated Stereo Rig**. 3rd European Conference on Computer Vision, v. 1, p. 567-576, 1994.
- [11] DI STEFANO, L.; MARCHIONNI, M.; MATTOCCIA S.; NERI, G. **A Fast Area-Based Stereo Matching Algorithm**. Image and Vision Computing, v. 22, p. 983-1005, 2004.
- [12] FAUGERAS, O. **Three-Dimensional Computer Vision: A Geometric Viewpoint**. MIT Press, 1993.
- [13] FELZENSZWALB, P. F.; HUTTENLOCHER, D. P. **Efficient Belief Propagation for Early Vision**. IEEE Conference on Computer Vision and Pattern Recognition, v. 1, p. 261-268, 2004.

- [14] GRIMSON, W. E. L. **From Images to Surfaces: A Computational Study of the Human Early Visual System**. MIT Press, 1981.
- [15] HADAMARD, J. S. **Sur les Problèmes aux Dérivées Partielles et leur Signification Physique**. Princeton University Bulletin, p. 49-52, 1902.
- [16] HIRSCHMULER, H. **Improvements in Real-Time Correlation-Based Stereo Vision**. IEEE Workshop on Stereo and Multi-Baseline Vision, p. 141-148, 2001.
- [17] JEONG H.; PARK, S.-C. **Trellis-Based Systolic Multi-Layer Stereo Matching**. IEEE Workshop on Signal Processing Systems, p. 257-262, 2003.
- [18] KANADE, T.; OKUTOMI, M. **A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment**. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 16, p. 920–932, 1994.
- [19] KLAUS, A.; SORMANN, M.; KARNER, K. **Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure**. 18th International Conference on Pattern Recognition, v. 3, p 15-18, 2006.
- [20] KOLMOGOROV, V.; ZABIH, R. **Computing Visual Correspondence with Occlusions via Graph Cuts**. 8th International Conference on Computer Vision, vol. 2, p. 508-515, 2001.
- [21] KOSCHAN, A. **What Is New in Computational Stereo Since 1989: A Survey on Current Stereo Papers**. Technical Report 93-22, University of Berlin, 1993.
- [22] LEWIS, J. P. **Fast Normalized Cross-Correlation**. Vision Interface, p. 120-123, 1995.
- [23] LONGUET-HIGGINS, H. C. **A Computer Algorithm for Reconstructing a Scene from Two Projections**. Nature, v. 293, p. 133-135, 1981.
- [24] LUONG, Q.-T.; FAUGERAS, O. D. **The Fundamental Matrix: Theory, Algorithms, and Stability Analysis**. International Journal of Computer Vision, v. 17, p. 43-75, 1996.
- [25] MARR, D. **Vision: A Computational Investigation into the Human Representation and Processing of Visual Information**. W. H. Freeman, 1982.
- [26] MARR, D.; POGGIO, T. **A Computational Theory of Human Stereo Vision**. Proceedings of the Royal Society of London, s. B, v. 204, p. 301-328, 1979.
- [27] MARR, D.; POGGIO, T. **Cooperative Computation of Stereo Disparity**. Science, 194, p. 283-287, 1976.
- [28] MULLIGAN J.; DANIILIDIS, K. **Predicting Disparity Windows for Real-Time Stereo**. 6th European Conference on Computer Vision, v. 1, p. 220-235, 2000.
- [29] NAKAMURA, Y.; MATSUURA, T.; SATOH, K.; OHTA, Y. **Occlusion Detectable Stereo – Occlusion Patterns in Camera Matrix**. IEEE Conference on Computer Vision and Pattern Recognition, p. 371–378, 1996.

- [30] OHTA Y.; KANADE. T. **Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming**. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 7, p. 139-154, 1985.
- [31] OKUTOMI M.; KANADE T. **A Multiple-Baseline Stereo**. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 15, p. 353-363, 1993.
- [32] OLIVEIRA, M. A. F. de; WAZLAWICK, R. S. **Linear Complexity Stereo Matching Based on Region Indexing**. 18th Brazilian Symposium on Computer Graphics and Image Processing, IEEE Computer Society Press, p. 181-188, 2005.
- [33] POGGIO, G. F.; POGGIO, T. **The Analysis of Stereopsis**. Annual Review of Neuroscience, v. 7, p. 379-412, 1984.
- [34] POGGIO, T.; TORRE, V. **Ill-posed Problems and Regularization Analysis in Early Vision**. DARPA Image Understanding Workshop, p. 257-263, 1984.
- [35] SCHARSTEIN D.; SZELISKI, R. **A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms**. International Journal of Computer Vision, vol 47, p. 7-42, 2002.
- [36] SCHARSTEIN D.; SZELISKI, R. **High-Accuracy Stereo Depth Maps Using Structured Light**. IEEE Conference on Computer Vision and Pattern Recognition, v. 1, p. 195-202, 2003.
- [37] SUN, C. **Fast Stereo Matching Using Rectangular Subregioning and 3D Maximum-Surface Techniques**. International Journal of Computer Vision, v. 47, p. 99-117, 2002.
- [38] SUN, J.; LI, Y.; KANG, S. B.; SHUM, H.-Y. **Symmetric Stereo Matching for Occlusion Handling**. IEEE Conference on Computer Vision and Pattern Recognition, v. 2, p. 399-406, 2005.
- [39] SZELISKI R.; ZABIH, R. **An Experimental Comparison of Stereo Algorithms**. International Workshop on Vision Algorithms, p. 1-19, 1999.
- [40] VEKSLER, O. **Dense Features for Semi-Dense Stereo Correspondence**. International Journal of Computer Vision, v. 47, p. 247-260, 2002.
- [41] VEKSLER, O. **Fast Variable Window for Stereo Correspondence Using Integral Images**. IEEE Conference on Computer Vision and Pattern Recognition, v. 1, p. 556-564, 2003.
- [42] YANG, Q.; WANG, L.; YANG, R.; WANG, S.; LIAO, M.; NISTÉR, D. **Real-Time Global Stereo Matching Using Hierarchical Belief Propagation**. 17th British Machine Vision Conference, v. 3, p. 989-998, 2006.
- [43] YANG, Q.; WANG, L.; YANG, R.; STEWÉNIUS, H.; NISTÉR, D. **Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling**. IEEE Conference on Computer Vision and Pattern Recognition, v. 2, p. 2347-2354, 2006.

Anexo 1 – Conjuntos de Dados



Conjuntos de dados padrão (de cima para baixo, “map”, “tsukuba”, “sawtooth”, “venus”, “cones” e “teddy”): (a) imagem esquerda; (b) imagem direita; (c) disparidades esperadas; (d) mapa de oclusões.

Anexo 2 – Plataforma Computacional

Configuração do computador pessoal utilizado em todos os testes:

tipo do processador	Pentium 4
CPUID	0F29h
frequência do processador	2,8 GHz
barramento do processador	800 MHz
tamanho da memória cache L2	512 KB
frequência da memória cache L2	2,8 GHz
tipo da memória principal	DDR, ECC, dual channel
tamanho da memória principal	2×512 MB
frequência da memória principal	400 MHz

Glossário

acurácia – Exatidão do resultado obtido em relação ao esperado, segundo metodologia de teste bem definida.

aspecto – Feição ou padrão geométrico pertencente a uma determinada classe e distinguível por seus atributos.

centro ótico – Centro de projeção da câmera. O ponto na cena, sua projeção no plano da câmera e o centro ótico são colineares.

correlação – Estimativa do grau de similaridade entre duas regiões, a qual pode ser baseada em valores de intensidade luminosa ou atributos de aspectos extraídos da imagem.

correlacionamento – Processo de associação de pontos correspondentes entre um par de imagens estéreo.

correspondência – Relação entre pontos de duas imagens. Disparidades entre pontos correspondentes são bidimensionais. A solução do problema da correspondência é interessante para detecção de continuidade temporal e estimativa de movimento.

correspondência estéreo – Relação entre pontos de duas imagens, havendo restrição epipolar. Disparidades entre pontos correspondentes são unidimensionais. A solução desse caso particular do problema da correspondência é interessante para estimativa de profundidade.

disparidade – Diferença de posição, em coordenadas da imagem, entre projeções de um mesmo ponto no espaço em ambas as imagens.

distância focal – Distância entre o plano de projeção e o centro ótico (comprimento do segmento de reta perpendicular ao plano de projeção que se estende dele até o centro ótico).

epipolo – Ponto sobre o plano de projeção da câmera onde todas as linhas epipolares se cruzam.

estereoscopia – Estimativa de profundidades com base na relação entre duas imagens obtidas da mesma cena, mas com ângulos levemente diferentes.

função de indexação – Função de avaliação de uma região da imagem para cálculo do valor discriminante correspondente. É chamada assim porque o valor discriminante é usado como índice no processo de indexação de regiões.

geometria epipolar – Geometria que descreve a restrição epipolar entre duas ou mais câmeras.

histograma – Número de ocorrências de cada intensidade em uma imagem ou de cada disparidade em um mapa de disparidades.

imagem – Função de domínio discreto e finito de intensidades luminosas sobre coordenadas cartesianas do plano de projeção.

interpolação – Processo de estimativa dos pontos sem valor definido no mapa de disparidades.

linha de base – Linha que contém os dois centros óticos de ambas as câmeras de um par estéreo.

linha epipolar – Intersecção entre o plano de projeção e o plano de triangulação, o qual contém a linha de base e o ponto triangulado. Linhas epipolares correspondentes são as interseções do mesmo plano de triangulação com ambos os plano de projeção.

mapa de disparidades – Conjunto de disparidades para cada ponto da imagem esquerda em relação a seu correspondente na direita.

par estéreo – Par de imagens obtidas ao mesmo tempo de uma cena, com pontos de vista levemente diferentes.

plano de projeção – Plano que contém a imagem, onde cada intensidade luminosa equivale à projeção de um ponto na cena.

profundidade – Distância entre a linha de base o ponto observado na cena (comprimento do segmento de reta perpendicular à linha de base que se estende dela até o ponto observado na cena).

região – Valores de intensidades próximos ou um aspecto bem definido em uma imagem.

regularização – Aplicação de uma regra de continuidade a um conjunto de pontos (como um mapa de disparidades), de modo a transformá-los segundo um determinado critério.

restrição epipolar – Projeções de um mesmo ponto na cena sobre ambos os planos de projeção sempre pertencem a linhas epipolares correspondentes.

retificação – Reprojeção dos pontos de ambas as imagens, onde os planos de projeção são coplanares, deslocando o epipolo para o infinito e tornando as linhas epipolares paralelas entre si e coincidentes com as linhas da imagem.

valor discriminante – Mapeamento que a função de indexação faz de uma determinada região. É usado como índice no processo de indexação de regiões.